(上級) 統計推理論

(2002 年度後期 講義ノート)

平成 14 年 9 月 5 日 (木) 版

谷﨑　久志
神戸大学・経済学部

目 次

1　事象と確率　1
- 1.1　事象　1
- 1.2　確率　1

2　確率変数と分布　2
- 2.1　1 次元の確率変数と分布　2
- 2.2　多次元の確率変数と分布　3
- 2.3　2.4 節のための数学の公式　4
  - 2.3.1　置換積分　4
  - 2.3.2　部分積分　5
  - 2.3.3　テーラー展開: 関数 $f(x)$ の近似　5
- 2.4　分布関数の持つ性質の証明 (いくつかの分布を例にとって)　5

3　平均値，分散　6
- 3.1　平均・分散の定義と公式　6
- 3.2　いくつかの分布の平均・分散　10

4　変数変換と和の分布 (連続型確率変数の場合のみ)　14
- 4.1　一変数の場合　14
- 4.2　二変数の場合　16

5　積率と積率母関数　19
- 5.1　積率母関数 (1 変数)　19
- 5.2　積率母関数 (多変数)　23

練習問題と解答 (1 章〜 5 章)　25

6　大数の法則と中心極限定理　34
- 6.1　Chebyshev の不等式　34
- 6.2　大数の (弱) 法則 (Convergence in probability)　34
- 6.3　中心極限定理　35

7　大数の強法則 (Almost sure convergence)　36

8　統計的推定　36
- 8.1　推定法と標本平均および標本分散の性質　36
  - 8.1.1　推定法　36
  - 8.1.2　標本，統計量，推定量　36
  - 8.1.3　母平均，母分散の推定　37
- 8.2　点推定法：最適性　38
- 8.3　推定量の求め方：最尤法, 積率法, 最小二乗法　42
  - 8.3.1　最尤法　42
  - 8.3.2　積率法 (モーメント法)　45
  - 8.3.3　最小二乗法　45

9 標本分布 45

9.1 正規母集団の場合 (標本平均，標本不偏分散の標本分布) 45

9.1.1 正規分布： 標本平均 $\overline{X}$ の標本分布 46

9.1.2 χ^2 (カイ自乗) 分布： 標本不偏分散 S^2 の標本分布 46

9.1.3 t 分布： 標本平均 $\overline{X}$ の標本分布 46

9.1.4 F 分布 47

9.2 その他の母集団の場合： 標本平均 $\overline{X}$ の標本分布 47

9.3 その他の母集団の場合： 母数の推定量 $\widehat{\theta}_n$ の標本分布 (一般化) 49

10 区間推定法 49

10.1 母平均 μ の区間推定 49

10.1.1 正規母集団の場合 (小標本，大標本共に) 49

10.1.2 その他の母集団の場合： 大標本 (n が大きいとき) 49

10.2 母分散 σ^2 の区間推定 (正規母集団) 50

10.3 母比率 p の区間推定 (ベルヌイ試行) 50

11 統計的検定 I 51

11.1 仮説検定の考え方 51

11.2 母平均 μ の検定 52

11.2.1 正規母集団の場合 (小標本，大標本共に) 52

11.2.2 その他の母集団の場合： 大標本 (n が大きいとき) 53

11.3 2 つの標本の母平均の差の検定 53

11.3.1 小標本 (n_1, n_2 が小さいとき) 53

11.3.2 その他の母集団：大標本 (n_1, n_2 が共に大きいとき) 54

11.4 母分散 σ^2 の検定で，正規母集団の場合 55

11.5 2 つの標本の母分散が等しいかどうかの検定で，正規母集団の場合 55

11.6 母比率 p の検定 (ベルヌイ試行) 56

12 統計的検定 II： 大標本検定 56

12.1 ワルド (Wald) 検定 56

12.2 尤度比検定 58

練習問題と解答 (6 章〜 12 章) 60

- この講義ノートは，
  http://ht.econ.kobe-u.ac.jp/~tanizaki/class
  からダウンロード可。

参考文献
- 『確率統計演習 1 確率』(国沢清典編，1966，培風館)
- 『確率統計演習 2 統計』(国沢清典編，1966，培風館)
- R.V. Hogg and A.T. Craig, 1995, *Introduction to Mathematical Statistics* (Fifth edition), Prentice Hall.

1 事象と確率

1.1 事象

試行，標本点，標本空間

試行： 考察の対象となる実験 (または，観測) を行うこと
標本点 ω： 試行によって得られる個々の結果
標本空間 Ω： 標本点全体の集合
例： サイコロ投げ：
サイコロ投げ 1 回の試行
標本点： 1, 2, 3, 4, 5, 6 の六つ
標本空間： $\Omega = \{1, 2, 3, 4, 5, 6\}$

事象とその演算

事象 A： 標本空間 Ω の部分集合
ω： 事象 A を構成する標本点の一つ
$\omega \in A$
例： サイコロ投げ：
サイコロ投げ 1 回の試行
$E = \{2, 4, 6\}$： 偶数の目が出る事象
$F = \{1, 2, 3\}$： 3 以下の目が出る事象
和事象：$E \cup F$： 事象 E と F のどちらか一方に属する標本点 ω の全体から成る集合
積事象：$E \cap F$： 事象 E と F のどちらにも属する標本点全体の集合
余事象：E^c： 事象 E に属さない標本点の集合
空事象：ϕ： 標本点を全然含まない事象
全事象：Ω： 全部を含む事象
排反：$E \cap F = \phi$ のとき，事象 E と F は互いに排反である

例：コイン投げ 3 回
表を H，裏を T とする。
標本点は次の 8 つ：
$\omega_1 = \{H, H, H\}$,
$\omega_2 = \{H, H, T\}$,
$\omega_3 = \{H, T, H\}$,
$\omega_4 = \{H, T, T\}$,
$\omega_5 = \{T, H, H\}$,
$\omega_6 = \{T, H, T\}$,
$\omega_7 = \{T, T, H\}$,
$\omega_8 = \{T, T, T\}$
標本空間： $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}$
2 回目が表であるという事象 E：
$E = \{\omega_1, \omega_2, \omega_5, \omega_6\}$
2 回表が出るという事象 F：
$F = \{\omega_2, \omega_3, \omega_5\}$
$E \cup F = \{\omega_1, \omega_2, \omega_3, \omega_5, \omega_6\}$
$E \cap F = \{\omega_2, \omega_5\}$
$E^c = \{\omega_3, \omega_4, \omega_7, \omega_8\}$
$F^c = \{\omega_1, \omega_4, \omega_6, \omega_7, \omega_8\}$
$(E \cup F)^c = \{\omega_4, \omega_7, \omega_8\}$
$E^c \cap F^c = \{\omega_4, \omega_7, \omega_8\}$
$(E \cup F)^c = E^c \cap F^c \Longrightarrow$ ド・モルガンの法則
$(E \cap F)^c = \{\omega_1, \omega_3, \omega_4, \omega_6, \omega_7, \omega_8\}$
$E^c \cup F^c = \{\omega_1, \omega_3, \omega_4, \omega_6, \omega_7, \omega_8\}$
$(E \cap F)^c = E^c \cup F^c \Longrightarrow$ ド・モルガンの法則

1.2 確率

事象 A の確率： $P(A)$
$0 \leq P(A) \leq 1$
$P(\Omega) = 1$, $P(\phi) = 0$
事象 A と B は互いに排反であるとき，$P(A \cup B) = P(A) + P(B)$
条件付き確率： 事象 B の条件のもとで事象 A の確率
$\Longrightarrow$
$P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
$P(A \cap B) = P(A|B)P(B) \Longrightarrow$ 乗法定理
事象 A と B は独立： $P(A \cap B) = P(A)P(B)$
公式：
$P(A^c) = 1 - P(A)$
$P(A \cup B) = P(A) + P(B) - P(A \cap B) \Longrightarrow$ 加法定理
$A \subset B$ のとき，$P(A) \leq P(B)$

2　確率変数と分布

2.1　1 次元の確率変数と分布

確率変数 X： 標本空間 Ω の上で定義された実数値関数 $X = X(\omega)$ を考える。
$X = X(\omega)$： 試行結果 (標本点) ω が定まると X の値が定まる。
$X(\omega)$ がある区間 I の中の値であるような標本点 ω の集合： $\{\omega; X(\omega) \in I\}$
$\{\omega; X(\omega) \in I\}$ を事象 $\{X \in I\}$ と書く。

離散型確率変数と確率分布：
確率変数 X の取りうる値を $a_1, a_2, \cdots$ とするとき,

$$P(X = a_i) = f(a_i), \quad i = 1, 2, \cdots$$

$f(a_i)$： X の確率分布
性質：

$$f(a_i) \geq 0, \quad i = 1, 2, \cdots$$
$$\sum_i f(a_i) = 1$$

ある集合 A について,

$$P(X \in A) = \sum_{a_i \in A} f(a_i)$$

となる。

連続型確率変数と確率密度関数：
ある区間 I について,

$$P(X \in I) = \int_I f(x) \mathrm{d}x$$

$f(x)$： X の確率密度関数
性質：

$$f(x) \geq 0,$$
$$\int_{-\infty}^{\infty} f(x) \mathrm{d}x = 1$$

また,

$$P(X = x) = \int_x^x f(t) \mathrm{d}t = 0,$$
$$P(X \in A) = \int_A f(x) \mathrm{d}x$$

分布関数：$P(X \leq x) = F(x)$

$F(x)$： X の分布関数
性質：
$x_1 < x_2$ のとき，$F(x_1) \leq F(x_2)$
$P(a < X \leq b) = F(b) - F(a)$
$F(-\infty) = 0$, $F(+\infty) = 1$

1. 離散型確率変数：
   $$F(x) = \sum_{a_i \leq x} f(a_i),$$
   $$F(a_i) - F(a_i - 0) = f(a_i)$$
2. 連続型確率変数：
   $$F(x) = \int_{-\infty}^{x} f(t) \mathrm{d}t,$$
   $$F'(x) = f(x)$$

重要な分布：

1. ベルヌイ分布：

   離散型確率変数 X の取りうる値は $0, 1$ のどちらかで, その確率分布は,
   $$P(X = k) = p^k (1-p)^{1-k}, \quad k = 0, 1$$
   $0 < p < 1$
2. 2 項分布：

   離散型確率変数 X の取りうる値が $0, 1, 2, \cdots, n$ で, その確率分布は,
   $$\begin{aligned} P(X = k) &= b(k; n, p) \\ &\equiv {}_nC_k p^k (1-p)^{n-k}, \\ &\qquad k = 0, 1, \cdots, n \end{aligned}$$
   $0 < p < 1$
3. ポアソン分布：

   離散型確率変数 X の取りうる値が $0, 1, 2, \cdots$ で，その確率分布は,
   $$\begin{aligned} P(X = k) &= p(k; \lambda) \\ &\equiv e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \cdots \end{aligned}$$
   $\lambda > 0$

$np = \lambda$ (一定) のもとで，$n \longrightarrow \infty$ のとき，

$$b(k;n,p) \ \longrightarrow \ p(k;\lambda)$$

4. 正規分布：

連続型確率変数 X の確率密度関数は，

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$X \sim N(\mu, \sigma^2)$

$N(0,1) \Longrightarrow$ 標準正規分布

5. 一様分布：

連続型確率変数 X の確率密度関数は，

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \text{ のとき} \\ 0, & \text{その他のとき} \end{cases}$$

6. 指数分布：

連続型確率変数 X の確率密度関数は，

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & 0 < x \text{ のとき} \\ 0, & \text{その他のとき} \end{cases}$$

$\lambda > 0$

$\lambda = \dfrac{1}{2}$ のとき，自由度 2 のカイ自乗分布に等しい。

7. χ^2 (カイ 2 乗) 分布 (自由度 n)：

連続型確率変数 X の確率密度関数は，

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\frac{n}{2})} 2^{-\frac{n}{2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x \ge 0 \text{ のとき} \\ 0, & x < 0 \text{ のとき} \end{cases}$$

$\Gamma(s) = \displaystyle\int_0^\infty u^{s-1} e^{-u} \mathrm{d}u \Longrightarrow$ ガンマ関数

$\Gamma(s+1) = s\Gamma(s)$, $\Gamma(1) = 1$, $\Gamma(\dfrac{1}{2}) = \sqrt{\pi}$

8. t 分布 (自由度 n)：

連続型確率変数 X の確率密度関数は，

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})} \frac{1}{\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

9. Cauchy 分布：

連続型確率変数 X の確率密度関数は，

$$f(x) = \frac{1}{\pi(1+x^2)}$$

自由度 1 の t 分布に等しい。

2.2 多次元の確率変数と分布

離散型確率変数 X と Y の取りうる値は $a_1, a_2, \cdots$ と $b_1, b_2, \cdots$ とする。

事象 $\{\omega; X(\omega) = a_i,$ かつ $Y(\omega) = b_j\}$ の確率は

$$P(X = a_i, Y = b_j) = h(a_i, b_j)$$

$h(a_i, b_j)$： X, Y の結合確率分布

性質：

$$h(a_i, b_j) \ge 0, \quad i, j = 1, 2, \cdots$$
$$\sum_{i,j} h(a_i, b_j) = 1$$

$f(a_i)$, $g(b_j)$ を次のように定義する。

$$f(a_i) = \sum_j h(a_i, b_j), \quad i = 1, 2, \cdots$$
$$g(b_j) = \sum_i h(a_i, b_j), \quad j = 1, 2, \cdots$$

$f(a_i)$, $g(b_j)$： X, Y の周辺確率分布

連続型確率変数 X と Y

ある領域 D について，事象 $\{\omega; \big(X(\omega), Y(\omega)\big) \in D\}$ の確率は

$$P\big((X,Y) \in D\big) = \iint_D h(x,y) \mathrm{d}x\mathrm{d}y$$

$h(x,y)$： X, Y の結合確率密度関数

性質：

$$h(x,y) \ge 0,$$
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) \mathrm{d}x\mathrm{d}y = 1$$

$f(x)$, $g(y)$ を次のように定義する。

$$f(x) = \int_{-\infty}^{\infty} h(x,y) \mathrm{d}y,$$
$$g(y) = \int_{-\infty}^{\infty} h(x,y) \mathrm{d}x,$$

$f(x)$, $g(y)$： X, Y の周辺確率密度関数

条件付き分布：

離散型：

$$P(X = a_i | Y = b_j) = f(a_i | b_j) \equiv \frac{h(a_i, b_j)}{g(b_j)}$$

$f(a_i|b_j)$： $Y=b_j$ を与えたもとで X の確率分布
性質：

$$f(a_i|b_j) \geq 0, \quad i=1,2,\cdots$$
$$\sum_i f(a_i|b_j) = 1$$

連続型：

$$f(x|y) = \frac{h(x,y)}{g(y)}$$

$f(x|y)$： $Y=y$ を与えたもとで X の確率密度関数
性質：

$$f(x|y) \geq 0,$$
$$\int_{-\infty}^{\infty} f(x|y) = 1$$

確率変数の独立性：
離散型： $h(a_i,b_j)=f(a_i)g(b_j)$ のとき，X と Y は独立となる。
連続型： $h(x,y)=f(x)g(y)$ のとき，X と Y は独立となる。

重要な分布：

1. 多項分布：

   離散型確率変数 $X_1,X_2,\cdots,X_r$ について，

$$P(X_1=k_1,X_2=k_2,\cdots,X_r=k_r) = \frac{n!}{k_1!k_2!\cdots k_r!}p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r}$$

   $k_1,k_2,\cdots,k_r$ は 0 以上の整数で，$\sum_{i=1}^{r} k_i = n$ を満たす。

   n は自然数

   $p_1,p_2,\cdots,p_r$ は正の定数で，$\sum_{i=1}^{r} p_i = 1$ を満たす。

2. 2 変数正規分布：
   連続型確率変数 X,Y の結合確率密度関数は

$$\begin{aligned}
&h(x,y) \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\
&\times \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}\right.\right. \\
&\qquad \left.\left.-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right) \\
&= \frac{1}{2\pi}\begin{vmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}^{-1/2} \\
&\times \exp\left(-\frac{1}{2}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}'\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}\right)
\end{aligned}$$

   μ_1, μ_2, σ_1, σ_2, ρ は定数で，$\sigma_1>0$, $\sigma_2>0$, $|\rho|<1$ とする。

   $\exp(x)$ は e^x と同じものであることに注意。

2.3　2.4 節のための数学の公式

2.3.1　置換積分

1 変数：　$f(x)$ について，$x=\psi(y)$ の置換積分を行う。

$$\int f(x)\mathrm{d}x = \int \psi'(y) f\big(\psi(y)\big)\mathrm{d}y$$

証明：

$$F(x) = \int f(x)\mathrm{d}x$$

$\Longrightarrow F'(x)=f(x)$
$F(x)=F\big(\psi(y)\big)$ を y について微分する。

$$\begin{aligned}
\frac{\mathrm{d}F\big(\psi(y)\big)}{\mathrm{d}y} &= \frac{\mathrm{d}F(x)}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}y} \\
&= f(x)\psi'(y) = f\big(\psi(y)\big)\psi'(y)
\end{aligned}$$

2 変数：　$f(x,y)$ について，$x=\psi_1(u,v)$, $y=\psi_2(u,v)$ のとき，

$$\begin{aligned}
&\int f(x,y)\mathrm{d}x\mathrm{d}y \\
&\qquad = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} f\big(\psi_1(u,v),\psi_2(u,v)\big)\mathrm{d}u\mathrm{d}v
\end{aligned}$$

(証明略)

$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とする。
$|A|=ad-bc$ を行列式の値と言う。

2.3.2 部分積分

$$\int f(x)g'(x)\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\mathrm{d}x$$

証明：
$f(x)g(x)$ の微分を考える。

$$\Big(f(x)g(x)\Big)' = f'(x)g(x) + f(x)g'(x)$$

両辺を積分すると,

$$\begin{aligned}&\int \Big(f(x)g(x)\Big)'\mathrm{d}x \\ &\quad = \int f'(x)g(x)\mathrm{d}x + \int f(x)g'(x)\mathrm{d}x\end{aligned}$$

となり,

$$f(x)g(x) = \int f'(x)g(x)\mathrm{d}x + \int f(x)g'(x)\mathrm{d}x$$

を得る。よって,

$$\int f(x)g'(x)\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\mathrm{d}x$$

2.3.3 テーラー展開: 関数 $f(x)$ の近似

$x = x_0$ の回りで $f(x)$ をテーラー展開する。

$$\begin{aligned}f(x) &\approx f(x_0) + f'(x_0)(x-x_0) \\ &\quad + \frac{1}{2!}f''(x_0)(x-x_0)^2 \\ &\quad + \frac{1}{3!}f'''(x_0)(x-x_0)^3 \\ &\quad + \cdots \\ &\quad = \sum_{n=0}^{\infty}\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n\end{aligned}$$

ただし，$f^{(n)}(x_0)$ は $f(x)$ を n 回微分して，$x = x_0$ で評価したものである。
$f^{(0)}(x_0) = f(x_0)$ と $0! = 1$ に注意。

2.4 分布関数の持つ性質の証明 (いくつかの分布を例にとって)

1. 2 項分布 $\displaystyle\sum_{k=0}^{n} b(k;n,p) = 1$ の証明：

$$\begin{aligned}&\sum_{k=0}^{n} b(k;n,p) \\ &= \sum_{k=0}^{n} {}_nC_k p^k (1-p)^{n-k} \\ &= \big(p + (1-p)\big)^n = 1 \text{ (2 項定理)}\end{aligned}$$

2. ポアソン分布 $\displaystyle\sum_{k=0}^{\infty} p(k;\lambda) = 1$ の証明：

$$\begin{aligned}\sum_{k=0}^{\infty} p(k;\lambda) &= \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} \\ &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} \\ &= e^{-\lambda}e^{\lambda} \\ &= 1\end{aligned}$$

$e^x = \displaystyle\sum_{k=0}^{\infty}\frac{x^k}{k!}$ に注意。

なぜなら，$f(x) = e^x$ としたとき，$f^{(k)}(x) = e^x$ となる。

テーラー展開の公式は,

$$f(x) = \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(x_0)(x-x_0)^k$$

なので，$x_0 = 0$ として，$x = 0$ の回りでテーラー展開すると,

$$\begin{aligned}f(x) &= \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)x^k \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}x^k \\ &= \sum_{k=0}^{\infty}\frac{x^k}{k!}\end{aligned}$$

を得る。

$f^{(n)}(0) = 1$ に注意。

3. 正規分布 $X \sim N(\mu, \sigma^2)$ の確率密度関数 $f(x)$ について，$\displaystyle\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$ の証明：

$$\begin{aligned}I &= \int_{-\infty}^{\infty} f(x)\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)\mathrm{d}u\end{aligned}$$

$u = \dfrac{x-\mu}{\sigma}$ として，置換積分を行う。

$\dfrac{\mathrm{d}x}{\mathrm{d}u} = \sigma$ に注意

$I = 1$ の証明は $I^2 = 1$ の証明を行えば十分

$$\begin{aligned}
I^2 &= (\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \mathrm{d}u) \\
&\times (\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}v^2\right) \mathrm{d}v) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(u^2+v^2)\right) \mathrm{d}u\mathrm{d}v \\
&= \frac{1}{2\pi}(\int_0^{2\pi} \mathrm{d}\theta)(\int_0^{\infty} \exp\left(-\frac{1}{2}r^2\right) r\mathrm{d}r) \\
&= \frac{1}{2\pi}(\int_0^{2\pi} \mathrm{d}\theta)(\int_0^{\infty} \exp(-s)\mathrm{d}s) \\
&= \frac{1}{2\pi} 2\pi [-\exp(-s)]_0^{\infty} \\
&= 1
\end{aligned}$$

$u = r\cos\theta$, $v = r\sin\theta$ として置換積分を行う。

$\begin{vmatrix} \dfrac{\partial u}{\partial r} & \dfrac{\partial u}{\partial \theta} \\ \dfrac{\partial v}{\partial r} & \dfrac{\partial v}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$

$0 < r < +\infty$, $0 < \theta < 2\pi$ となることに注意

さらに，$s = \dfrac{1}{2}r^2$ と置換積分される。

このように，$I^2 = 1$ が得られ，$f(x) \geq 0$ なので，$I = 1$ を得る。

4. 指数分布に従う X の確率密度関数 $f(x)$ について，$\displaystyle\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$ の証明：

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\mathrm{d}x &= \int_0^{\infty} \lambda e^{-\lambda x}\mathrm{d}x \\
&= [-e^{-\lambda x}]_0^{\infty} \\
&= 1
\end{aligned}$$

5. 一様分布に従う X の確率密度関数 $f(x)$ について，$\displaystyle\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$ の証明：

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\mathrm{d}x &= \int_a^b \frac{1}{b-a}\mathrm{d}x \\
&= [\frac{1}{b-a}x]_a^b \\
&= 1
\end{aligned}$$

6. X, Y は 2 変数正規分布に従うとき，X の周辺確率密度関数は？

連続型確率変数 X, Y の結合確率密度関数は

$$\begin{aligned}
h(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\
&\quad \times \exp\Biggl(-\frac{1}{2(1-\rho^2)}(\frac{(x-\mu_1)^2}{\sigma_1^2} \\
&\qquad -2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \\
&\qquad +\frac{(y-\mu_2)^2}{\sigma_2^2})\Biggr)
\end{aligned}$$

$$\begin{aligned}
&-\frac{1}{2(1-\rho^2)}(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \\
&+\frac{(y-\mu_2)^2}{\sigma_2^2}) \\
=& -\frac{1}{2(1-\rho^2)}(\frac{y-\mu_2}{\sigma_2} - \rho\frac{x-\mu_1}{\sigma_1})^2 \\
&-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2} \\
=& -\frac{1}{2(1-\rho^2)\sigma_2^2}\bigl((y-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\bigr)^2 \\
&-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}
\end{aligned}$$

$$\begin{aligned}
f(x) &= \int_{-\infty}^{\infty} h(x,y)\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right) \\
&\times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma^2}} \exp\Biggl(-\frac{1}{2(1-\rho^2)\sigma_2^2} \\
&\qquad \times\bigl((y-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\bigr)^2\Biggr)\mathrm{d}y
\end{aligned}$$

積分の部分は，$N\bigl(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), (1-\rho^2)\sigma_2^2\bigr)$ に対応し，積分値は 1 になる。

3 平均値，分散

3.1 平均・分散の定義と公式

1 変数： 確率変数 X のある関数：$g(X)$

定義：

$g(X)$ の期待値 $\mathrm{E}\big(g(X)\big)$：

$$\mathrm{E}\big(g(X)\big)=\begin{cases}\displaystyle\sum_i g(x_i)p_i=\sum_i g(x_i)f(x_i), & \\ \qquad\qquad\qquad\text{離散型確率変数} & \\ \displaystyle\int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x, & \\ \qquad\qquad\qquad\text{連続型確率変数} & \end{cases}$$

1. 確率変数 X の平均 $\mathrm{E}(X)$

   $\Longrightarrow X$ の期待値, $g(X)=X$

$$\mathrm{E}(X)=\begin{cases}\displaystyle\sum_i x_i f(x_i), & \text{離散型確率変数}\\ \displaystyle\int_{-\infty}^{\infty} xf(x)\mathrm{d}x, & \text{連続型確率変数}\end{cases}$$
$$=\mu,\quad (\text{または，}\mu_x)$$

2. 確率変数 X の分散 $\mathrm{V}(X)$

   $\Longrightarrow (X-\mu)^2$ の期待値, $g(X)=(X-\mu)^2$

$$\begin{aligned}\mathrm{V}(X)&=\mathrm{E}\big((X-\mu)^2\big)\\&=\begin{cases}\displaystyle\sum_i (x_i-\mu)^2 f(x_i), & \\ \qquad\qquad\text{離散型確率変数} & \\ \displaystyle\int_{-\infty}^{\infty}(x-\mu)^2 f(x)\mathrm{d}x, & \\ \qquad\qquad\text{連続型確率変数} & \end{cases}\\&=\sigma^2,\quad(\text{または，}\sigma_x^2)\end{aligned}$$

   確率変数 X の分散 $\mathrm{V}(X)$

   $\Longrightarrow X$ の確率分布の確率関数 (離散型の場合)，または，確率密度関数 (連続型の場合) の範囲が広ければ, $\mathrm{V}(X)$ は大きい。

いくつかの公式：

1. a, b を定数とする。

   定理： $\mathrm{E}(aX+b)=a\mathrm{E}(X)+b$

   証明：

   X が離散型確率変数の場合,

$$\begin{aligned}\mathrm{E}(aX+b)&=\sum_i(ax_i+b)f(x_i)\\&=a\sum_i x_i f(x_i)+b\sum_i f(x_i)\\&=a\mathrm{E}(X)+b\end{aligned}$$

   途中で，$\displaystyle\sum_i f(x_i)=1$ に注意

   X が連続型確率変数の場合,

$$\begin{aligned}\mathrm{E}(aX+b)&=\int_{-\infty}^{\infty}(ax+b)f(x)\mathrm{d}x\\&=a\int_{-\infty}^{\infty}xf(x)\mathrm{d}x+b\int_{-\infty}^{\infty}f(x)\mathrm{d}x\\&=a\mathrm{E}(X)+b\end{aligned}$$

   途中で，$\displaystyle\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$ に注意

2. **定理：** $\mathrm{V}(X)=\mathrm{E}(X^2)-\mu^2$

   証明：

$$\begin{aligned}\mathrm{V}(X)&=\mathrm{E}\big((X-\mu)^2\big)\\&=\mathrm{E}(X^2-2\mu X-\mu^2)\\&=\mathrm{E}(X^2)-2\mu\mathrm{E}(X)+\mu^2\\&=\mathrm{E}(X^2)-\mu^2\end{aligned}$$

   途中で，$\mu=\mathrm{E}(X)$ に注意

3. a, b を定数とする。

   定理： $\mathrm{V}(aX+b)=a^2\mathrm{V}(X)$

   証明：

   $\mathrm{E}(aX+b)=a\mu+b$ に注意して,

$$\begin{aligned}\mathrm{V}(aX+b)&=\mathrm{E}\Big(\big((aX+b)-\mathrm{E}(aX+b)\big)^2\Big)\\&=\mathrm{E}\Big((aX-a\mu)^2\Big)\\&=\mathrm{E}\big(a^2(X-\mu)^2\big)\\&=a^2\mathrm{E}\big((X-\mu)^2\big)\\&=a^2\mathrm{V}(X)\end{aligned}$$

   を得る。

4. **定理：** 確率変数 X について，$\mathrm{E}(X)=\mu$, $\mathrm{V}(X)=\sigma^2$ とする。
$Z=\dfrac{X-\mu}{\sigma}$ を定義する。
このとき，$\mathrm{E}(Z)=0$, $\mathrm{V}(Z)=1$ となる。

証明：

$$\mathrm{E}(Z)=\mathrm{E}\left(\frac{X-\mu}{\sigma}\right)=\frac{\mathrm{E}(X)-\mu}{\sigma}=0$$

$$\mathrm{V}(Z)=\mathrm{V}\left(\frac{1}{\sigma}X-\frac{\mu}{\sigma}\right)=\frac{1}{\sigma^2}\mathrm{V}(X)=1$$

2 変数： 確率変数 X,Y のある関数： $g(X,Y)$
定義：
$g(X,Y)$ の期待値 $\mathrm{E}\big(g(X,Y)\big)$：

$$\mathrm{E}\big(g(X,Y)\big)=\begin{cases}\displaystyle\sum_i\sum_j g(x_i,y_j)f(x_i,y_j), & \\ \qquad\qquad \text{離散型確率変数} & \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y, & \\ \qquad\qquad \text{連続型確率変数} & \end{cases}$$

1. 確率変数 X の平均 $\mathrm{E}(X)$

   $\Longrightarrow$ X の期待値, $g(X,Y)=X$

$$\mathrm{E}(X)=\begin{cases}\displaystyle\sum_i\sum_j x_i f(x_i,y_j), & \\ \qquad\qquad \text{離散型確率変数} & \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\mathrm{d}x\mathrm{d}y, & \\ \qquad\qquad \text{連続型確率変数} & \end{cases}$$
$$=\mu_x$$

2. 確率変数 X の分散 $\mathrm{V}(X)$

   $\Longrightarrow$ $(X-\mu_x)^2$ の期待値, $g(X,Y)=(X-\mu_x)^2$

$$\mathrm{V}(X)=\mathrm{E}\big((X-\mu_x)^2\big)$$

$$=\begin{cases}\displaystyle\sum_i\sum_j (x_i-\mu_x)^2 f(x_i,y_j), & \\ \qquad\qquad \text{離散型確率変数} & \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)^2 f(x,y)\mathrm{d}x\mathrm{d}y, & \\ \qquad\qquad \text{連続型確率変数} & \end{cases}$$
$$=\sigma_x^2$$

3. 確率変数 X,Y の共分散 $\mathrm{Cov}(X,Y)$

   $\Longrightarrow$ $(X-\mu_x)(Y-\mu_y)$ の期待値, $g(X,Y)=(X-\mu_x)(Y-\mu_y)$

$$\begin{aligned}&\mathrm{Cov}(X,Y)\\ &=\mathrm{E}\big((X-\mu_x)(Y-\mu_y)\big)\\ &=\begin{cases}\displaystyle\sum_i\sum_j (x_i-\mu_x)(y_j-\mu_y) f(x_i,y_j), & \\ \qquad\qquad \text{離散型確率変数} & \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y) f(x,y)\mathrm{d}x\mathrm{d}y, & \\ \qquad\qquad \text{連続型確率変数} & \end{cases}\end{aligned}$$

いくつかの公式：

1. 確率変数 X,Y について，

   定理： $\mathrm{E}(X+Y)=\mathrm{E}(X)+\mathrm{E}(Y)$

   証明：

$$\begin{aligned}\mathrm{E}(X+Y)&=\sum_i\sum_j(x_i+y_j)f(x_i,y_j)\\ &=\sum_i\sum_j x_i f(x_i,y_j)\\ &\quad+\sum_i\sum_j y_j f(x_i,y_j)\\ &=\mathrm{E}(X)+\mathrm{E}(Y)\end{aligned}$$

2. 確率変数 X と Y が独立のとき，

   定理： $\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$

   証明：

$$\mathrm{E}(XY)=\sum_i\sum_j x_i y_j f(x_i,y_j)$$

$$
\begin{aligned}
&= \sum_i \sum_j x_i y_j f(x_i) h(y_j) \\
&= \sum_i x_i f(x_i) \sum_j y_j h(y_j) \\
&= \mathrm{E}(X)\mathrm{E}(Y)
\end{aligned}
$$

3. 確率変数 X, Y について，

定理： $\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$

証明：

$$
\begin{aligned}
&\mathrm{Cov}(X,Y) \\
&= \sum_i \sum_j (x_i - \mu_x)(y_j - \mu_y) p_{ij} \\
&= \sum_i \sum_j (x_i y_j - \mu_x y_j - \mu_y x_i + \mu_x \mu_y) p_{ij} \\
&= \sum_i \sum_j x_i y_j f(x_i, y_j) \\
&- \mu_x \sum_i \sum_j y_j f(x_i, y_j) \\
&- \mu_y \sum_i \sum_j x_i f(x_i, y_j) \\
&+ \mu_x \mu_y \sum_i \sum_j f(x_i, y_j) \\
&= \sum_i \sum_j x_i y_j f(x_i, y_j) - \mu_x \mu_y - \mu_y \mu_x + \mu_x \mu_y \\
&= \sum_i \sum_j x_i y_j f(x_i, y_j) - \mu_x \mu_y \\
&= \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)
\end{aligned}
$$

より，一般的な証明：

$$
\begin{aligned}
&\mathrm{Cov}(X,Y) \\
&= \mathrm{E}\big((X - \mu_x)(Y - \mu_y)\big) \\
&= \mathrm{E}\big(XY - \mu_x Y - \mu_y X + \mu_x \mu_y\big) \\
&= \mathrm{E}(XY) - \mathrm{E}(\mu_x Y) - \mathrm{E}(\mu_y X) + \mu_x \mu_y \\
&= \mathrm{E}(XY) - \mu_x \mathrm{E}(Y) - \mu_y \mathrm{E}(X) + \mu_x \mu_y \\
&= \mathrm{E}(XY) - \mu_x \mu_y - \mu_y \mu_x + \mu_x \mu_y \\
&= \mathrm{E}(XY) - \mu_x \mu_y \\
&= \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)
\end{aligned}
$$

4. 確率変数 X と Y が独立のとき，

$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$

となるので，

$\mathrm{Cov}(X,Y) = 0$

を得る。

5. 相関係数 ρ_{xy}：

$$
\begin{aligned}
\rho_{xy} &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)}\sqrt{\mathrm{V}(Y)}} \\
&= \frac{\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y}
\end{aligned}
$$

6. 確率変数 X と Y が独立のとき，

$\mathrm{Cov}(X,Y) = 0$

となるので，

$\rho_{xy} = 0$

を得る。

7. 確率変数 X, Y について，

$\mathrm{V}(X \pm Y) = \mathrm{V}(X) \pm 2\mathrm{Cov}(X,Y) + \mathrm{V}(Y)$

証明：

$$
\begin{aligned}
&\mathrm{V}(X \pm Y) \\
&= \mathrm{E}\Big(\big((X \pm Y) - \mathrm{E}(X \pm Y)\big)^2\Big) \\
&= \mathrm{E}\Big(\big((X - \mu_x) \pm (Y - \mu_y)\big)^2\Big) \\
&= \mathrm{E}\big((X - \mu_x)^2 \pm 2(X - \mu_x)(Y - \mu_y) \\
&\qquad +(Y - \mu_y)^2\big) \\
&= \mathrm{E}\big((X - \mu_x)^2\big) \\
&\pm 2\mathrm{E}\big((X - \mu_x)(Y - \mu_y)\big) \\
&+ \mathrm{E}\big((Y - \mu_y)^2\big) \\
&= \mathrm{V}(X) \pm 2\mathrm{Cov}(X,Y) + \mathrm{V}(Y)
\end{aligned}
$$

8. $-1 \leq \rho_{xy} \leq 1$

証明：

次のような t に関する式を考える。$f(t) = \mathrm{V}(Xt - Y)$

分散なので，必ずゼロ以上となる。よって，すべての t について，$f(t) \geq 0$ となるための条件を求めればよい。t に関する 2 次方程式の判別式がゼロ以下となる条件

を求める。$\mathrm{V}(Xt-Y)=\mathrm{V}(Xt)-2\mathrm{Cov}(Xt,Y)+\mathrm{V}(Y)$
$=t^2\mathrm{V}(X)-2t\mathrm{Cov}(X,Y)+\mathrm{V}(Y)$

$\frac{D}{2}=\big(\mathrm{Cov}(X,Y)\big)^2-\mathrm{V}(X)\mathrm{V}(Y)\leq 0$

$\frac{\big(\mathrm{Cov}(X,Y)\big)^2}{\mathrm{V}(X)\mathrm{V}(Y)}\leq 1$

$-1\leq\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)}\sqrt{\mathrm{V}(Y)}}\leq 1$

$-1\leq\rho_{xy}\leq 1$

ρ_{xy} が 1 に近いほど，正の相関が強くなる。

ρ_{xy} が -1 に近いほど，負の相関が強くなる。

9. 確率変数 X と Y が独立のとき，

定理： $\mathrm{V}(X+Y)=\mathrm{V}(X)+\mathrm{V}(Y)$

証明：

$\mathrm{V}(X+Y)=\mathrm{V}(X)+2\mathrm{Cov}(X,Y)+\mathrm{V}(Y)$

確率変数 X と Y が独立のとき，

$\mathrm{Cov}(X,Y)=0$

なので，

$\mathrm{V}(X+Y)=\mathrm{V}(X)+\mathrm{V}(Y)$

を得る。

10. n 個の確率変数 $X_1,\, X_2,\, \cdots,\, X_n$ について：

$\mathrm{E}(X_i)=\mu_i$ とするとき，

$\mathrm{E}(\sum_i X_i)=\sum_i\mathrm{E}(X_i)=\sum_i\mu_i$

$\mathrm{V}(\sum_i X_i)\equiv\mathrm{E}\Big(\sum_i(X_i-\mu_i)\Big)^2$

$=\mathrm{E}\Big(\sum_i(X_i-\mu_i)\Big)\Big(\sum_j(X_j-\mu_j)\Big)$

$=\mathrm{E}\Big(\sum_i\sum_j(X_i-\mu_i)(X_j-\mu_j)\Big)$

$=\sum_i\sum_j\mathrm{E}\Big((X_i-\mu_i)(X_j-\mu_j)\Big)$

$=\sum_i\sum_j\mathrm{Cov}(X_i,X_j)$

11. n 個の確率変数 $X_1,\, X_2,\, \cdots,\, X_n$ は互いに独立で同じ平均 μ，分散 σ^2 を持つとする。すなわち，すべての $i=1,2,\cdots,n$ について，

$$\mathrm{E}(X_i)=\mu,\quad \mathrm{V}(X_i)=\sigma^2$$

を仮定する。

さらに，算術平均 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ を考える。

このとき，

定理： $\mathrm{E}(\overline{X})=\mu,\, \mathrm{V}(\overline{X})=\frac{\sigma^2}{n}$

が成り立つ。

証明：

$$\begin{aligned}\mathrm{E}(\overline{X})&=\mathrm{E}(\sum_i\frac{X_i}{n})\\&=\sum_i\mathrm{E}(\frac{X_i}{n})\\&=\sum_i\frac{1}{n}\mathrm{E}(X_i)\\&=\sum_i\frac{1}{n}\mu\\&=\mu\end{aligned}$$

$$\begin{aligned}\mathrm{V}(\overline{X})&=\mathrm{V}(\sum_i\frac{X_i}{n})\\&=\sum_i\mathrm{V}(\frac{X_i}{n})\\&=\sum_i\frac{1}{n^2}\mathrm{V}(X_i)\\&=\sum_i\frac{1}{n}^2\sigma^2\\&=\frac{\sigma^2}{n}\end{aligned}$$

3.2 いくつかの分布の平均・分散

ベルヌイ分布の平均と分散： ベルヌイ分布：

$$f(x)=p^x(1-p)^{1-x}\qquad x=0,1$$

$\mathrm{E}(X)=p,\, \mathrm{V}(X)=p(1-p)$
証明：
平均：

$$\begin{aligned}\mathrm{E}(X)&=\sum_x xf(x)\\&=\sum_{x=0}^{1}xp^x(1-p)^{1-x}\\&=p\end{aligned}$$

分散：
$\mu = \mathrm{E}(X)$ のとき，$\mathrm{V}(X) = \mathrm{E}(X^2) - \mu^2$ により，$\mathrm{E}(X^2)$ を求める。

$$\begin{aligned}\mathrm{E}(X^2) &= \sum_x x^2 f(x)\\ &= \sum_{x=0}^{1} x^2 p^x (1-p)^{1-x}\\ &= p^2\end{aligned}$$

よって，

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}(X^2) - \mu^2\\ &= p - p^2\\ &= p(1-p)\end{aligned}$$

2 項分布の平均と分散： 2 項分布：

$$\begin{aligned}f(x) &= {}_nC_x p^x (1-p)^{n-x}\\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \cdots, n,\end{aligned}$$

$\mathrm{E}(X) = np$, $\mathrm{V}(X) = np(1-p)$
証明：
平均：

$$\begin{aligned}\mathrm{E}(X) &= \sum_x x f(x)\\ &= \sum_x x\, {}_nC_x p^x (1-p)^{n-x}\\ &= \sum_x x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}\\ &= \sum_x \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x}\\ &= np \sum_x \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x}\\ &= np \sum_{x'} \frac{n'!}{x'!(n'-x')!} p^{x'} (1-p)^{n'-x'}\\ &= np \sum_{x'} {}_{n'}C_{x'} p^{x'} (1-p)^{n'-x'}\\ &= np\end{aligned}$$

ただし，$n' = n-1$, $x' = x-1$ と定義される。
確率関数の性質より，

$$\sum_x {}_nC_x p^x (1-p)^{n-x} = 1$$

を得ることに注意。
分散：
$\mathrm{V}(X) = \mathrm{E}(X^2) - \mu^2$ により，$\mathrm{E}(X^2)$ を求める。
$X^2 = X(X-1) + X$ を利用する。
$\mathrm{E}(X^2) = \mathrm{E}\big(X(X-1)\big) + \mathrm{E}(X)$
したがって，
$\mathrm{V}(X) = \mathrm{E}\big(X(X-1)\big) + \mu - \mu^2$ となる。
右辺第 1 項を求める。

$$\begin{aligned}&\mathrm{E}\big(X(X-1)\big)\\ &= \sum_x x(x-1) f(x)\\ &= \sum_x x(x-1)\, {}_nC_x p^x (1-p)^{n-x}\\ &= \sum_x x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}\\ &= \sum_x \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x}\\ &= n(n-1)p^2 \sum_x \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x}\\ &= n(n-1)p^2 \sum_{x'} \frac{n'!}{x'!(n'-x')!} p^{x'} (1-p)^{n'-x'}\\ &= n(n-1)p^2 \sum_{x'} {}_{n'}C_{x'} p^{x'} (1-p)^{n'-x'}\\ &= n(n-1)p^2\end{aligned}$$

途中で，$n' = n-2$, $x' = x-2$ と定義されている。
まとめると，

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}(X^2) - \mu^2\\ &= \mathrm{E}\big(X(X-1)\big) + \mu - \mu^2\\ &= n(n-1)p^2 + np - n^2p^2\\ &= -np^2 + np\\ &= np(1-p)\end{aligned}$$

ポアソン分布の平均と分散： ポアソン分布：

$$f(x) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots, n,$$

$\mathrm{E}(X) = \lambda$, $\mathrm{V}(X) = \lambda$
証明：
平均：

$$\mathrm{E}(X) = \sum_k k f(k)$$

$$
\begin{aligned}
&= \sum_k k e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \sum_k \lambda e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda \sum_{k'} e^{-\lambda} \frac{\lambda^{k'}}{k'!} \\
&= \lambda
\end{aligned}
$$

ただし，$k' = k - 1$ と定義される。
確率関数の性質より，

$$
\sum_k e^{-\lambda} \frac{\lambda^k}{k!} = 1,
$$

を得ることに注意。
分散：
$\mathrm{V}(X) = \mathrm{E}(X^2) - \mu^2$ により，$\mathrm{E}(X^2)$ を求める。
$X^2 = X(X-1) + X$ を利用する。
$\mathrm{E}(X^2) = \mathrm{E}\big(X(X-1)\big) + \mathrm{E}(X)$
したがって，
$\mathrm{V}(X) = \mathrm{E}\big(X(X-1)\big) + \mu - \mu^2$ となる。
右辺第 1 項を求める。

$$
\begin{aligned}
&\mathrm{E}\big(X(X-1)\big) \\
&= \sum_k k(k-1) f(k) \\
&= \sum_k k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \sum_k \lambda^2 e^{-\lambda} \frac{\lambda^{k-2}}{(k-2)!} \\
&= \lambda^2 \sum_{k'} e^{-\lambda} \frac{\lambda^{k'}}{k'!}
\end{aligned}
$$

途中で，$k' = k - 2$ と定義されている。
まとめると，

$$
\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}(X^2) - \mu^2 \\
&= \mathrm{E}\big(X(X-1)\big) + \mu - \mu^2 \\
&= \lambda^2 + \lambda - \lambda^2 \\
&= \lambda
\end{aligned}
$$

正規分布の平均と分散：　正規分布：$X \sim N(\mu, \sigma^2)$

$$
f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}
$$

$\mathrm{E}(X) = \mu$, $\mathrm{V}(X) = \sigma^2$
証明：
平均：

$$
\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \big((x-\mu) + \mu\big) f(x) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} (x-\mu) f(x) \mathrm{d}x + \mu \int_{-\infty}^{\infty} f(x) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \mathrm{d}x + \mu \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \left[-\sigma^2 e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \right]_{-\infty}^{\infty} + \mu \\
&= \mu
\end{aligned}
$$

確率密度関数の性質にから，

$$
\begin{aligned}
&\int_{-\infty}^{\infty} f(x) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \mathrm{d}x \\
&= 1
\end{aligned}
$$

となることに注意。

合成関数の微分：
$y = h\big(g(x)\big) \Longrightarrow y = h(u)$, $u = g(x)$

$$
\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\mathrm{d}y}{\mathrm{d}u} \frac{\mathrm{d}u}{\mathrm{d}x} \\
&= h'(u) g'(x) \\
&= h\big(g(x)\big) g'(x)
\end{aligned}
$$

上の計算では，
$h(u) = -\sigma^2 e^u$, $g(x) = -\dfrac{1}{2\sigma^2}(x-\mu)^2$
とすればよい。

ロピタルの定理：
ある関数 $g(x)$, $f(x)$ について，

$$
\lim_{x \to \infty} \frac{f'(x)}{g'(x)} = A
$$

ならば，

$$
\lim_{x \to \infty} \frac{f(x)}{g(x)} = A
$$

となる。
例：
$\lim_{x\to\infty}\frac{x}{e^x}=\lim_{x\to\infty}\frac{1}{e^x}=0$
$\frac{\mathrm{d}e^x}{\mathrm{d}x}=e^x$ に注意。
分散：
$\mathrm{V}(X)=\mathrm{E}(X-\mu)^2$

$$\begin{aligned}\mathrm{V}(X)&=\int_{-\infty}^{\infty}(x-\mu)^2f(x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}(x-\mu)^2\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x\\&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}(x-\mu)\frac{\mathrm{d}\left(-\sigma^2e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\right)}{\mathrm{d}x}\mathrm{d}x\\&=\frac{1}{\sqrt{2\pi\sigma^2}}\left[(x-\mu)\left(-\sigma^2e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\right)\right]_{-\infty}^{\infty}\\&+\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\sigma^2e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x\\&=\sigma^2\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x\\&=\sigma^2\end{aligned}$$

部分積分：
$\int_a^b h(x)g'(x)\mathrm{d}x=\Big[h(x)g(x)\Big]_a^b-\int_a^b h'(x)g(x)\mathrm{d}x$ を利用。
$h(x)=x-\mu$, $g(x)=-\sigma^2e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ とする。

指数分布の平均と分散： 指数分布：

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & 0<x \text{ のとき}\\ 0, & \text{その他のとき}\end{cases}$$

$\mathrm{E}(X)=\frac{1}{\lambda}$, $\mathrm{V}(X)=\frac{1}{\lambda^2}$
証明：
平均：

$$\begin{aligned}\mathrm{E}(X)&=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x\\&=\int_0^{\infty}x\lambda e^{-\lambda x}\mathrm{d}x\\&=\int_0^{\infty}x\frac{\mathrm{d}(-e^{-\lambda x})}{\mathrm{d}x}\mathrm{d}x\\&=\left[x(-e^{-\lambda x})\right]_0^{\infty}+\int_0^{\infty}e^{-\lambda x}\mathrm{d}x\\&=\frac{1}{\lambda}\int_0^{\infty}\lambda e^{-\lambda x}\mathrm{d}x\\&=\frac{1}{\lambda}\end{aligned}$$

確率密度関数の性質から,

$$\begin{aligned}\int_{-\infty}^{\infty}f(x)\mathrm{d}x&=\int_0^{\infty}\lambda e^{-\lambda x}\mathrm{d}x\\&=[-e^{-\lambda x}]_0^{\infty}\\&=1\end{aligned}$$

に注意。
分散：
$\mathrm{V}(X)=\mathrm{E}(X^2)-\mu^2$ により，$\mathrm{E}(X^2)$ を求める。

$$\begin{aligned}\mathrm{E}(X^2)&=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x\\&=\int_0^{\infty}x^2\lambda e^{-\lambda x}\mathrm{d}x\\&=\int_0^{\infty}x^2\frac{\mathrm{d}(-e^{-\lambda x})}{\mathrm{d}x}\mathrm{d}x\\&=\left[x^2(-e^{-\lambda x})\right]_0^{\infty}+2\int_0^{\infty}xe^{-\lambda x}\mathrm{d}x\\&=\frac{2}{\lambda}\int_0^{\infty}x\lambda e^{-\lambda x}\mathrm{d}x\\&=\frac{2}{\lambda}\mathrm{E}(X)\\&=\frac{2}{\lambda^2}\end{aligned}$$

$\mathrm{V}(X)=\mathrm{E}(X^2)-\mu^2$
$=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}$
$=\frac{1}{\lambda^2}$

一様分布の平均と分散： 一様分布：

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b \text{ のとき}\\ 0, & \text{その他のとき}\end{cases}$$

$\mathrm{E}(X)=\frac{b-a}{2}$, $\mathrm{V}(X)=\frac{(b-a)^2}{12}$
証明：
平均：

$$\begin{aligned}\mathrm{E}(X)&=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x\\&=\int_a^b x\frac{1}{b-a}\mathrm{d}x\\&=\frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b\\&=\frac{a+b}{2}\end{aligned}$$

分散：
$\mathrm{V}(X)=\mathrm{E}(X^2)-\mu^2$ により，$\mathrm{E}(X^2)$ を求める。

$$\begin{aligned}\mathrm{E}(X^2)&=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x\\&=\int_a^b x^2\frac{1}{b-a}\mathrm{d}x\\&=\frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b\\&=\frac{a^2+ab+b^2}{3}\end{aligned}$$

$\mathrm{V}(X)=\dfrac{a^2+ab+b^2}{3}-\Big(\dfrac{a+b}{2}\Big)^2$
$=\dfrac{(b-a)^2}{12}$

カイ二乗分布の平均と分散： カイ二乗分布：

$$f(x)=\begin{cases}\dfrac{1}{\Gamma(\frac{n}{2})}2^{-\frac{n}{2}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, & x\geq 0\text{ のとき}\\ 0, & x<0\text{ のとき}\end{cases}$$

$\mathrm{E}(X)=n$, $\mathrm{V}(X)=2n$
証明：
平均：

$$\begin{aligned}\mathrm{E}(X)&=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x\\&=\int_0^{\infty}x\frac{1}{\Gamma(\frac{n}{2})}2^{-\frac{n}{2}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=\frac{2^{-\frac{n}{2}}}{2^{-\frac{n+2}{2}}}\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n}{2})}\\&\quad\times\int_0^{\infty}\frac{1}{\Gamma(\frac{n+2}{2})}2^{-\frac{n+2}{2}}x^{\frac{n+2}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=2\frac{n}{2}\int_0^{\infty}\frac{1}{\Gamma(\frac{n'}{2})}2^{-\frac{n'}{2}}x^{\frac{n'}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=n\end{aligned}$$

$\Gamma(s)=\displaystyle\int_0^{\infty}u^{s-1}e^{-u}\mathrm{d}u\Longrightarrow$ ガンマ関数 $\Gamma(s+1)=s\Gamma(s)$, $\Gamma(1)=1$, $\Gamma(\frac{1}{2})=\sqrt{\pi}$ に注意
また，$n'=n+2$ を使い，確率密度関数の性質から，

$$\int_{-\infty}^{\infty}f(x)\mathrm{d}x=\int_0^{\infty}\frac{1}{\Gamma(\frac{n}{2})}2^{-\frac{n}{2}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x=1$$

に注意。
分散：
$\mathrm{V}(X)=\mathrm{E}(X^2)-\mu^2$ により，$\mathrm{E}(X^2)$ を求める。

$$\begin{aligned}\mathrm{E}(X^2)&=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x\\&=\int_0^{\infty}x^2\frac{1}{\Gamma(\frac{n}{2})}2^{-\frac{n}{2}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=\int_0^{\infty}\frac{1}{\Gamma(\frac{n}{2})}2^{-\frac{n}{2}}x^{\frac{n+4}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=\frac{2^{-\frac{n}{2}}}{2^{-\frac{n+4}{2}}}\frac{\Gamma(\frac{n+4}{2})}{\Gamma(\frac{n}{2})}\\&\quad\times\int_0^{\infty}\frac{1}{\Gamma(\frac{n+4}{2})}2^{-\frac{n+4}{2}}x^{\frac{n+4}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=4(\frac{n+2}{2}\frac{n}{2})\int_0^{\infty}\frac{1}{\Gamma(\frac{n'}{2})}2^{-\frac{n'}{2}}x^{\frac{n'}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\&=n(n+2)\end{aligned}$$

$n'=n+4$ を使う。
$\mathrm{V}(X)=n(n+2)-n^2=2n$

4 変数変換と和の分布 (連続型確率変数の場合のみ)

4.1 一変数の場合

$Y=\psi^{-1}(X)$ の分布： 連続型確率変数 X の確率密度関数 $f(x)$
$X=\psi(Y)$ の一対一変換のとき，
Y の確率密度関数 $g(y)$ は，

$$g(y)=|\psi'(y)|f\big(\psi(y)\big)$$

証明：
確率変数 X の分布関数を $F(x)$，確率密度関数を $f(x)$ とする。
(すなわち，$F(x)=P(X\leq x)$, $f(x)=F'(x)$ である。)
$Y=h(X)$ とする。
$X=\psi(Y)$ のとき，Y の分布を求める。
すなわち，$h^{-1}(Y)=\psi(Y)$ となる。
Y の分布関数を $G(y)$，確率密度関数を $g(y)$ とする。
$\psi'(X)>0$ の場合：

$$\begin{aligned}G(y)&=P(Y\leq y)\\&=P\Big(h(X)\leq y\Big)\\&=P\Big(X\leq h^{-1}(y)\Big)\end{aligned}$$

$$
\begin{aligned}
&= P\Big(X \le \psi(y)\Big) \\
&= F\Big(\psi(y)\Big)
\end{aligned}
$$

なので，

$$
\begin{aligned}
g(y) &= G'(y) \\
&= \psi'(y)F'\Big(\psi(y)\Big) \\
&= \psi'(y)f\Big(\psi(y)\Big)
\end{aligned}
$$

を得る。

$\psi'(X) < 0$ の場合：

$$
\begin{aligned}
G(y) &= P(Y \le y) \\
&= P\Big(h(X) \le y\Big) \\
&= P\Big(X \ge h^{-1}(y)\Big) \\
&= P\Big(X \ge \psi(y)\Big) \\
&= 1 - P\Big(X < \psi(y)\Big) \\
&= 1 - F\Big(\psi(y)\Big)
\end{aligned}
$$

なので，

$$
\begin{aligned}
g(y) &= G'(y) \\
&= -\psi'(y)F'\Big(\psi(y)\Big) \\
&= -\psi'(y)f\Big(\psi(y)\Big)
\end{aligned}
$$

を得る。$-\psi'(y) > 0$ に注意

したがって，まとめると，

$$
g(y) = |\psi'(y)|f\Big(\psi(y)\Big)
$$

を得る。$\Longrightarrow$ 変数変換

$Y = X^2$ の分布について： 確率変数 X の分布関数を $F(x)$，確率密度関数を $f(x)$ とする。

確率変数 Y の分布関数を $G(y)$，確率密度関数を $g(y)$ とする。

$Y = X^2$ の分布関数 $G(y)$ は，

$$
\begin{aligned}
G(y) &= P(Y \le y) \\
&= P(X^2 \le y) \\
&= P(-\sqrt{y} \le X \le \sqrt{y}) \\
&= F(\sqrt{y}) - F(-\sqrt{y})
\end{aligned}
$$

したがって，

$$
\begin{aligned}
g(y) &= G'(y) \\
&= \frac{1}{2\sqrt{y}}\Big(f(\sqrt{y}) + f(-\sqrt{y})\Big)
\end{aligned}
$$

例： $\chi^2(1)$ 分布： $X \sim N(0,1)$ とするとき，$Y = X^2 \sim \chi^2(1)$ となる。

証明：

X の分布関数とその微分:

$$
F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \mathrm{d}u
$$

$$
f(x) = F'(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)
$$

Y の確率密度関数 $g(y)$ は，$y > 0$ について，

$$
\begin{aligned}
g(y) &= G'(y) \\
&= \frac{1}{2\sqrt{y}}\Big(F'(\sqrt{y}) + F'(-\sqrt{y})\Big) \\
&= \frac{1}{\sqrt{2\pi}\sqrt{y}} \exp\left(-\frac{1}{2}y\right) \\
&= \frac{1}{\Gamma(\frac{1}{2})2^{\frac{1}{2}}} y^{\frac{1}{2}-1} \exp\left(-\frac{1}{2}y\right)
\end{aligned}
$$

これは $Y \sim \chi^2(1)$ を意味する。

$\Gamma(\frac{1}{2}) = \sqrt{\pi}$ に注意

$Y \sim \chi^2(n)$ のとき，Y の確率密度関数は，

$$
f(y) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} \exp(-\frac{y}{2})
$$

に注意

例： $N(0,1)$ 分布： $X \sim N(\mu, \sigma^2)$ とするとき，$Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$ となる。

証明：

X の分布関数:

$$
f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)
$$

$X = \psi(Y) = \sigma Y + \mu$ なので，Y の密度関数は，

$$
\begin{aligned}
g(y) &= |\psi'(y)|f\Big(\psi(y)\Big) \\
&= |\sigma|\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}y^2\right) \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right)
\end{aligned}
$$

となる。これは，$N(0,1)$ に一致する。

4.2 二変数の場合

$X=\psi_1(U,V)$，$Y=\psi_2(U,V)$ のとき，(U,V) の分布： 確率変数 X, Y の結合密度関数 $f(x,y)$ について，$X=\psi_1(U,V)$, $Y=\psi_2(U,V)$ のとき，確率変数 U, V の結合密度関数 $g(u,v)$

$$g(u,v)=\begin{vmatrix}\dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v}\\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v}\end{vmatrix} f\big(\psi_1(u,v),\psi_2(u,v)\big)$$

(証明略)

U の周辺密度関数 $h(u)$:

$$h(u)=\int g(u,v)\mathrm{d}v$$

V の周辺密度関数 $p(v)$:

$$p(v)=\int g(u,v)\mathrm{d}u$$

例： 正規分布： X, Y は互いに独立な確率変数でそれぞれ正規分布 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ に従うとき，$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$ となる。

証明：

X, Y の確率密度関数

$$f(x)=\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left(-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right)$$
$$g(y)=\frac{1}{\sqrt{2\pi\sigma_2^2}}\exp\left(-\frac{1}{2\sigma_2^2}(y-\mu_2)^2\right)$$

X, Y の結合確率密度関数は，X, Y は互いに独立な確率変数なので，

$$\begin{aligned}&h(x,y)\\&=f(x)g(y)\\&=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2\sigma_1^2}(x-\mu_1)^2-\frac{1}{2\sigma_2^2}(y-\mu_2)^2\right)\end{aligned}$$

$U=X+Y$, $V=Y$ として，U, V の結合確率密度関数を求める。$X=U-V$, $Y=V$ なので，

$$\begin{pmatrix}\dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v}\\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v}\end{pmatrix}=\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$$

となるので，U, V の結合確率密度関数 $s(u,v)$ は，変数変換により，

$$\begin{aligned}&s(u,v)\\&=h(u-v,v)\begin{vmatrix}1 & -1\\ 0 & 1\end{vmatrix}\\&=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2\sigma_1^2}(u-v-\mu_1)^2-\frac{1}{2\sigma_2^2}(v-\mu_2)^2\right)\end{aligned}$$

U の周辺確率密度関数 $p(u)$ を求める。

$$\begin{aligned}&p(u)\\
&=\int s(u,v)\mathrm{d}v\\
&=\int\frac{1}{2\pi\sigma_1\sigma_2}\exp\Big(-\frac{1}{2\sigma_1^2}(u-v-\mu_1)^2\\
&\qquad-\frac{1}{2\sigma_2^2}(v-\mu_2)^2\Big)\mathrm{d}v\\
&=\int\frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2\sigma_1^2}\big((v-\mu_2)-(u-\mu_1-\mu_2)\big)^2\right.\\
&\qquad\left.-\frac{1}{2\sigma_2^2}(v-\mu_2)^2\right)\mathrm{d}v\\
&=\int\frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2/(1/\sigma_1^2+1/\sigma_2^2)}\big((v-\mu_2)\right.\\
&\qquad-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}(u-\mu_1-\mu_2)\big)^2\\
&\qquad\left.-\frac{1}{2(\sigma_1^2+\sigma_2^2)}(u-\mu_1-\mu_2)^2\right)\mathrm{d}v\\
&=\int\frac{1}{\sqrt{2\pi/(1/\sigma_1^2+1/\sigma_2^2)}}\\
&\quad\times\exp\left(-\frac{1}{2/(1/\sigma_1^2+1/\sigma_2^2)}\big((v-\mu_2)\right.\\
&\qquad\left.-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}(u-\mu_1-\mu_2)\big)^2\right)\\
&\times\frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\left(-\frac{1}{2(\sigma_1^2+\sigma_2^2)}(u-\mu_1-\mu_2)^2\right)\mathrm{d}v\\
&=\int\frac{1}{\sqrt{2\pi/(1/\sigma_1^2+1/\sigma_2^2)}}\\
&\quad\times\exp\left(-\frac{1}{2/(1/\sigma_1^2+1/\sigma_2^2)}\big((v-\mu_2)\right.\\
&\qquad\left.-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}(u-\mu_1-\mu_2)\big)^2\right)\mathrm{d}v\\
&\times\frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\left(-\frac{1}{2(\sigma_1^2+\sigma_2^2)}(u-\mu_1-\mu_2)^2\right)\\
&=\frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\left(-\frac{1}{2(\sigma_1^2+\sigma_2^2)}(u-\mu_1-\mu_2)^2\right)
\end{aligned}$$

例： χ^2 分布： X, Y は互いに独立で，$X \sim \chi^2(n), Y \sim \chi^2(m)$ とするとき，$U = X + Y \sim \chi^2(n+m)$ となる。

証明：

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}\exp(-\frac{x}{2}), \quad x > 0$$
$$g(y) = \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m}{2}-1}\exp(-\frac{y}{2}), \quad y > 0$$

X, Y の結合確率密度関数は，X, Y は互いに独立な確率変数なので，

$$\begin{aligned}
&h(x,y) \\
&= f(x)g(y) \\
&= \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}\exp(-\frac{x}{2}) \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m}{2}-1}\exp(-\frac{y}{2}) \\
&= Cx^{\frac{n}{2}-1}y^{\frac{m}{2}-1}\exp(-\frac{x+y}{2})
\end{aligned}$$

ただし，$C = \dfrac{1}{2^{\frac{n+m}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}$ とする。

$U = X + Y, V = Y$ として，変数変換を行う。

$X = U - V, Y = V$ なので，

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

となるので，U, V の結合確率密度関数 $s(u,v)$ は，変数変換により，

$$\begin{aligned}
&s(u,v) \\
&= h(u-v,v)\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} \\
&= C(u-v)^{\frac{n}{2}-1}v^{\frac{m}{2}-1}\exp(-\frac{u}{2})
\end{aligned}$$

U の周辺確率密度関数は，

$$\begin{aligned}
&p(u) \\
&= \int s(u,v)\mathrm{d}v \\
&= C\exp(-\frac{u}{2})\int_0^\infty (u-v)^{\frac{n}{2}-1}v^{\frac{m}{2}-1}\mathrm{d}v \\
&= C\exp(-\frac{u}{2})\int_0^\infty (u-uw)^{\frac{n}{2}-1}(uw)^{\frac{m}{2}-1}u\mathrm{d}w \\
&= Cu^{\frac{n+m}{2}-1}\exp(-\frac{u}{2})\int_0^\infty (1-w)^{\frac{n}{2}-1}w^{\frac{m}{2}-1}\mathrm{d}w \\
&= CB(\frac{n}{2},\frac{m}{2})u^{\frac{n+m}{2}-1}\exp(-\frac{u}{2}) \\
&= \frac{1}{2^{\frac{n+m}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}\frac{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}{\Gamma(\frac{n+m}{2})}u^{\frac{n+m}{2}-1}\exp(-\frac{u}{2}) \\
&= \frac{1}{2^{\frac{n+m}{2}}\Gamma(\frac{n+m}{2})}u^{\frac{n+m}{2}-1}\exp(-\frac{u}{2})
\end{aligned}$$

$w = \dfrac{v}{u}$，すなわち，$v = uw$ として置換積分 ($\dfrac{\mathrm{d}v}{\mathrm{d}w} = u$ に注意)

ベータ関数 $B(n,m)$ は

$$\begin{aligned}
B(n,m) &= \int_0^\infty (1-x)^{n-1}x^{m-1}\mathrm{d}x \\
&= \frac{\Gamma(n)\Gamma(m)}{\Gamma(n+m)}
\end{aligned}$$

に注意

例： t 分布： X, Y は互いに独立で，$X \sim N(0,1), Y \sim \chi^2(n)$ とするとき，$U = \dfrac{X}{\sqrt{Y/n}} \sim t(n)$ となる。

U の密度関数 $f(u)$ は，

$$f(u) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})}\frac{1}{\sqrt{n\pi}}\left(1 + \frac{u^2}{n}\right)^{-\frac{n+1}{2}}$$

となる。

証明：

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2), \quad -\infty < x < \infty$$
$$g(y) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}y^{\frac{n}{2}-1}\exp(-\frac{y}{2}), \quad y > 0$$

X, Y の結合確率密度関数は，X, Y は互いに独立な確率変数なので，

$$\begin{aligned}
&h(x,y) \\
&= f(x)g(y) \\
&= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}y^{\frac{n}{2}-1}\exp(-\frac{y}{2}) \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}y^{\frac{n}{2}-1}\exp(-\frac{y}{2}-\frac{1}{2}x^2)
\end{aligned}$$

$U = \dfrac{X}{\sqrt{Y/n}}, V = Y$ として，変数変換を行う。

$X = U\sqrt{\dfrac{V}{n}}, Y = V$ なので，

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{v}{n}} & \frac{u}{2\sqrt{nv}} \\ 0 & 1 \end{pmatrix}$$

となるので，U, V の結合確率密度関数 $s(u,v)$ は，変数変換により，

$$
\begin{aligned}
&s(u,v)\\
&= h(u\sqrt{\frac{v}{n}},v)\begin{vmatrix}\sqrt{\frac{v}{n}} & \frac{u}{2\sqrt{nv}}\\ 0 & 1\end{vmatrix}\\
&= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}\frac{u^2v}{n})\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}v^{\frac{n}{2}-1}\exp(-\frac{v}{2})\sqrt{\frac{v}{n}}\\
&= Cv^{\frac{n-1}{2}}\exp\left(-\frac{v}{2}(1+\frac{u^2}{n})\right)
\end{aligned}
$$

ただし，$C=\frac{1}{\sqrt{\pi}}\frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})}\frac{1}{\sqrt{n}}$ とする。
U の周辺確率密度関数は，

$$
\begin{aligned}
&p(u)\\
&= \int s(u,v)\mathrm{d}v\\
&= C\int v^{\frac{n-1}{2}}\exp\left(-\frac{v}{2}(1+\frac{u^2}{n})\right)\mathrm{d}v\\
&= C\int\left(w(1+\frac{u^2}{n})^{-1}\right)^{\frac{n-1}{2}}\exp(-\frac{1}{2}w)\left(1+\frac{u^2}{n}\right)^{-1}\mathrm{d}w\\
&= C\left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}}\int w^{\frac{n+1}{2}-1}\exp(-\frac{1}{2}w)\mathrm{d}w\\
&= C\left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}}2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})\\
&\times\int\frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})}w^{\frac{n+1}{2}-1}\exp(-\frac{1}{2}w)\mathrm{d}w\\
&= C\left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}}2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})\\
&= \frac{1}{\sqrt{\pi}}\frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})}\frac{1}{\sqrt{n}}2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})\left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}}\\
&= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})}\frac{1}{\sqrt{n\pi}}\left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}}
\end{aligned}
$$

$w=v(1+\frac{u^2}{n})$ として置換積分。
$f(w)=\frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})}w^{\frac{n+1}{2}-1}\exp(-\frac{1}{2}w)$ は $\chi^2(n+1)$ の密度関数に注意。

例：Cauchy 分布： X, Y は互いに独立で，$X\sim N(0,1)$, $Y\sim N(0,1)$ とするとき，$U=\frac{X}{Y}$ は Cauchy 分布となる。
U の密度関数 $f(u)$ は，

$$
f(u)=\frac{1}{\pi(1+u^2)}
$$

となる。
証明：

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2),\quad -\infty<x<\infty\\
g(y) &= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2),\quad -\infty<y<\infty
\end{aligned}
$$

X, Y の結合確率密度関数は，X, Y は互いに独立な確率変数なので，

$$
\begin{aligned}
&h(x,y)\\
&= f(x)g(y)\\
&= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2)\\
&= \frac{1}{2\pi}\exp\bigl(-\frac{1}{2}(x^2+y^2)\bigr)
\end{aligned}
$$

$u=\frac{x}{y}$, $v=y$ として，変数変換を行う。
$x=uv$, $y=v$ なので，

$$
\begin{pmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{pmatrix}=\begin{pmatrix}v & u\\ 0 & 1\end{pmatrix}
$$

となるので，U, V の結合確率密度関数 $s(u,v)$ は，変数変換により，

$$
\begin{aligned}
&s(u,v)\\
&= h(uv,v)\begin{vmatrix}v & u\\ 0 & 1\end{vmatrix}\\
&= \frac{1}{2\pi}\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)|v|
\end{aligned}
$$

U の周辺確率密度関数は，

$$
\begin{aligned}
&p(u)\\
&= \int s(u,v)\mathrm{d}v\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}|v|\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)\mathrm{d}v\\
&= \frac{1}{\pi}\int_0^{\infty}v\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)\mathrm{d}v
\end{aligned}
$$

$$= \frac{1}{\pi}\left[-\frac{1}{1+u^2}\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)\right]_{v=0}^{\infty}$$
$$= \frac{1}{\pi(1+u^2)}$$

5 積率と積率母関数

5.1 積率母関数 (1 変数)

確率変数 X について，
積率： $\mu'_n = \mathrm{E}(X^n) \Longrightarrow$ 原点のまわりの n 次の積率
積率母関数： $\phi(\theta) = \mathrm{E}(e^{\theta X})$

$$\phi(\theta) = \mathrm{E}(e^{\theta X}) = \begin{cases} \displaystyle\sum_i e^{\theta a_i} f(a_i), & \text{離散型} \\ \displaystyle\int_{-\infty}^{\infty} e^{\theta x} f(x)\mathrm{d}x, & \text{連続型} \end{cases}$$

性質：

1. $\phi^{(n)}(0) = \mu'_n \equiv \mathrm{E}(X^n)$

   証明：

   $\phi(\theta)$ を $\theta = 0$ のまわりで，テーラー展開を行う。

   $$\begin{aligned} &\phi(\theta) \\ &= \mathrm{E}\bigl(1 + \frac{X}{1!}\theta + \frac{X^2}{2!}\theta^2 + \cdots + \frac{X^n}{n!}\theta^n + \cdots\bigr) \\ &= 1 + \frac{\mathrm{E}(X)}{1!}\theta + \frac{\mathrm{E}(X^2)}{2!}\theta^2 + \cdots + \frac{\mathrm{E}(X^n)}{n!}\theta^n + \cdots \\ &= 1 + \frac{\mu'_1}{1!}\theta + \frac{\mu'_2}{2!}\theta^2 + \cdots + \frac{\mu'_n}{n!}\theta^n + \cdots \end{aligned}$$

   したがって，

   $$\phi^{(n)}(\theta) = \mu'_n + \frac{\mu'_{n+1}}{1!}\theta + \frac{\mu'_{n+2}}{2!}\theta^2 + \cdots$$

   より，$\phi^{(n)}(0) = \mu'_n \equiv \mathrm{E}(X^n)$

   注)

   関数 $f(x)$ の $x = x_0$ の回りでテーラー展開

   $$f(x) \approx \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k$$

   ただし，$f^{(k)}(x_0)$ は $f(x)$ の k 回微分を $x = x_0$ で評価したものとする。

2. 確率変数 X の積率母関数と確率変数 Y の積率母関数は一致するとき，確率変数 X の分布関数と確率変数 Y の分布関数も一致する。

3. 互いに独立な確率変数 $X_1, X_2, \cdots, X_n$ の積率母関数を $\phi_1(\theta), \phi_2(\theta), \cdots, \phi_n(\theta)$ とするとき，$X_1+X_2+\cdots+X_n$ の積率母関数は $\phi_1(\theta)\phi_2(\theta)\cdots\phi_n(\theta)$ となる。

   証明：

   $Y = X_1 + X_2 + \cdots + X_n$ として，Y の積率母関数 $\phi_y(\theta)$ は

   $$\begin{aligned} \phi_y(\theta) &= \mathrm{E}(e^{\theta Y}) \\ &= \mathrm{E}(e^{\theta(X_1+X_2+\cdots+X_n)}) \\ &= \mathrm{E}(e^{\theta X_1})\mathrm{E}(e^{\theta X_2})\cdots\mathrm{E}(e^{\theta X_n}) \\ &= \phi_1(\theta)\phi_2(\theta)\cdots\phi_n(\theta) \end{aligned}$$

4. 互いに独立な確率変数 X_1, X_2, $\cdots$, X_n が同一の分布に従い，その積率母関数を $\phi(\theta)$ とするとき，$X_1 + X_2 + \cdots + X_n$ の積率母関数は $\bigl(\phi(\theta)\bigr)^n$ となる。

ベルヌイ分布の積率母関数： $\phi(\theta) = pe^\theta + (1-p)$

ベルヌイ分布

$$f(x) = p^x(1-p)^{1-x}, \quad x = 0, 1$$

積率母関数

$$\begin{aligned} \phi(\theta) &= \sum_{x=0}^{1} e^{\theta x} f(x) \\ &= \sum_{x=0}^{1} e^{\theta x} p^x (1-p)^{1-x} \\ &= e^\theta p + 1 - p \end{aligned}$$

1. 平均：

   $\mathrm{E}(X) = \phi'(0)$

   $\phi(\theta) = e^\theta p + 1 - p$,

   $\phi'(\theta) = pe^\theta$ なので，

   $$\begin{aligned} \mathrm{E}(X) &= \phi'(0) \\ &= p \end{aligned}$$

2. 分散：

$\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

$\mathrm{E}(X^2) = \phi''(0)$

$\phi''(\theta) = pe^\theta$ なので，

$$\begin{aligned} \mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\ &= \phi''(0) - \big(\phi'(0)\big)^2 \\ &= p - p^2 \\ &= p(1-p) \end{aligned}$$

二項分布の積率母関数： $\phi(\theta) = \Big(pe^\theta + (1-p)\Big)^n$

二項分布

$$f(x) = {}_nC_x p^x (1-p)^{n-x}, \quad x = 0, 1, \cdots, n$$

積率母関数

$$\begin{aligned} \phi(\theta) &= \sum_{x=0}^{n} e^{\theta x} f(x) \\ &= \sum_{x=0}^{n} e^{\theta x} {}_nC_x p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} {}_nC_x (e^\theta p)^x (1-p)^{n-x} \\ &= (e^\theta p + 1 - p)^n \end{aligned}$$

(二項定理より)

1. 平均：

$\mathrm{E}(X) = \phi'(0)$

$\phi(\theta) = (e^\theta p + 1 - p)^n$,

$\phi'(\theta) = npe^\theta(e^\theta p + 1 - p)^{n-1}$ なので，

$$\begin{aligned} \mathrm{E}(X) &= \phi'(0) \\ &= np \end{aligned}$$

2. 分散：

$\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

$\mathrm{E}(X^2) = \phi''(0)$

$\phi''(\theta) = npe^\theta(e^\theta p + 1 - p)^{n-1} + n(n-1)p^2 e^{2\theta}(e^\theta p + 1 - p)^{n-2}$ なので，

$$\begin{aligned} \mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\ &= \phi''(0) - \big(\phi'(0)\big)^2 \\ &= np + n(n-1)p^2 - (np)^2 \\ &= np(1-p) \end{aligned}$$

ポアソン分布の積率母関数： $\phi(\theta) = \exp\Big(\lambda(e^\theta - 1)\Big)$

ポアソン分布

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0, 1, \cdots$$

積率母関数

$$\begin{aligned} \phi(\theta) &= \sum_{x=0}^{\infty} e^{\theta x} f(x) \\ &= \sum_{x=0}^{\infty} e^{\theta x} e^{-\lambda}\frac{\lambda^x}{x!} \\ &= \sum_{x=0}^{\infty} e^{-\lambda}\frac{(e^\theta\lambda)^x}{x!} \\ &= \exp(-\lambda)\exp(e^\theta\lambda) \\ &= \exp\Big(\lambda(e^\theta - 1)\Big) \end{aligned}$$

$\displaystyle\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^\lambda$ に注意

1. 平均：

$\mathrm{E}(X) = \phi'(0)$

$\phi(\theta) = \exp\Big(\lambda(e^\theta - 1)\Big)$,

$\phi'(\theta) = \lambda e^\theta \exp\Big(\lambda(e^\theta - 1)\Big)$ なので，

$$\begin{aligned} \mathrm{E}(X) &= \phi'(0) \\ &= \lambda \end{aligned}$$

2. 分散：

$\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

$\mathrm{E}(X^2) = \phi''(0)$

$\phi''(\theta) = (1 + \lambda e^\theta)\lambda e^\theta \exp\Big(\lambda(e^\theta - 1)\Big)$ なので，

$$\begin{aligned} \mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\ &= \phi''(0) - \big(\phi'(0)\big)^2 \\ &= (1+\lambda)\lambda - \lambda^2 \\ &= \lambda \end{aligned}$$

正規分布の積率母関数： $\phi(\theta) = \exp\Big(\mu\theta + \dfrac{1}{2}\sigma^2\theta^2\Big)$

正規分布

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$$

積率母関数

$$\begin{aligned}
\phi(\theta) &= \int_{-\infty}^{\infty} e^{\theta x} f(x)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2+\theta x}\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x^2-2(\mu+\sigma^2\theta)x+\mu^2\right)}\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x-(\mu+\sigma^2\theta)\right)^2+(\mu\theta+\frac{1}{2}\sigma^2\theta^2)}\mathrm{d}x \\
&= e^{\mu\theta+\frac{1}{2}\sigma^2\theta^2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x-(\mu+\sigma^2\theta)\right)^2}\mathrm{d}x \\
&= \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right)
\end{aligned}$$

積分のところは，$N(\mu+\sigma^2\theta, \sigma^2)$ の確率密度関数に注意

1. 平均：

   $\mathrm{E}(X) = \phi'(0)$

   $\phi(\theta) = \exp\Big(\mu\theta + \dfrac{1}{2}\sigma^2\theta^2\Big)$,

   $\phi'(\theta) = (\mu+\sigma^2\theta)\exp\Big(\mu\theta + \dfrac{1}{2}\sigma^2\theta^2\Big)$ なので，

   $$\begin{aligned}
   \mathrm{E}(X) &= \phi'(0) \\
   &= \mu
   \end{aligned}$$

2. 分散：

   $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

   $\mathrm{E}(X^2) = \phi''(0)$

   $\phi''(\theta) = \sigma^2\exp\Big(\mu\theta + \dfrac{1}{2}\sigma^2\theta^2\Big) + (\mu+\sigma^2\theta)^2\exp\Big(\mu\theta + \dfrac{1}{2}\sigma^2\theta^2\Big)$ なので，

   $$\begin{aligned}
   \mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\
   &= \phi''(0) - \big(\phi'(0)\big)^2 \\
   &= (\sigma^2+\mu^2) - \mu^2 \\
   &= \sigma^2
   \end{aligned}$$

一様分布の積率母関数： $\phi(\theta) = \dfrac{e^{b\theta} - e^{a\theta}}{\theta(b-a)}$

一様分布

$$f(x) = \frac{1}{b-a}, \quad a < x < b$$

積率母関数

$$\begin{aligned}
\phi(\theta) &= \int_{-\infty}^{\infty} e^{\theta x} f(x)\mathrm{d}x \\
&= \int_{a}^{b} e^{\theta x} \frac{1}{b-a}\mathrm{d}x \\
&= \left[\frac{e^{\theta x}}{\theta(b-a)}\right]_a^b \\
&= \frac{e^{\theta b} - e^{\theta a}}{\theta(b-a)}
\end{aligned}$$

1. 平均：

   $\mathrm{E}(X) = \phi'(0)$

   $\phi(\theta) = \dfrac{e^{\theta b} - e^{\theta a}}{\theta(b-a)}$,

   $\phi'(\theta) = \dfrac{be^{\theta b} - ae^{\theta a}}{\theta(b-a)} - \dfrac{e^{\theta b} - e^{\theta a}}{\theta^2(b-a)}$ なので，

   $$\begin{aligned}
   \mathrm{E}(X) &= \phi'(0) \\
   &= (a+b) - \frac{a+b}{2} \\
   &= \frac{a+b}{2}
   \end{aligned}$$

2. 分散：

   $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

   $\mathrm{E}(X^2) = \phi''(0)$

   $\phi''(\theta) = \dfrac{b^2e^{\theta b} - a^2e^{\theta a}}{\theta(b-a)} - 2\dfrac{be^{\theta b} - ae^{\theta a}}{\theta^2(b-a)} + 2\dfrac{e^{\theta b} - e^{\theta a}}{\theta^3(b-a)}$

   なので，

   $$\begin{aligned}
   \mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\
   &= \phi''(0) - \big(\phi'(0)\big)^2 \\
   &= \Big((b^2+ab+a^2) - 2\frac{b^2+ab+a^2}{2} \\
   &\qquad + 2\frac{b^2+ab+a^2}{6}\Big) - \Big(\frac{a+b}{2}\Big)^2 \\
   &= \frac{(b-a)^2}{12}
   \end{aligned}$$

指数分布の積率母関数： $\phi(\theta) = \dfrac{\lambda}{\lambda - \theta}$

指数分布

$$f(x) = \lambda e^{-\lambda x}, \quad 0 < x$$

積率母関数

$$\begin{aligned}\phi(\theta) &= \int_{-\infty}^{\infty} e^{\theta x} f(x)\mathrm{d}x \\ &= \int_{0}^{\infty} e^{\theta x} \lambda e^{-\lambda x}\mathrm{d}x \\ &= \frac{\lambda}{\lambda - \theta}\int_{0}^{\infty} (\lambda - \theta) e^{-(\lambda-\theta)x}\mathrm{d}x \\ &= \frac{\lambda}{\lambda - \theta}\end{aligned}$$

積分のところは，パラメータ $\lambda - \theta$ の指数分布に注意

1. 平均：

   $\mathrm{E}(X) = \phi'(0)$

   $\phi(\theta) = \dfrac{\lambda}{\lambda - \theta}$,

   $\phi'(\theta) = \dfrac{\lambda}{(\lambda - \theta)^2}$ なので，

   $$\begin{aligned}\mathrm{E}(X) &= \phi'(0) \\ &= \frac{1}{\lambda}\end{aligned}$$

2. 分散：

   $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

   $\mathrm{E}(X^2) = \phi''(0)$

   $\phi''(\theta) = 2\dfrac{\lambda}{(\lambda - \theta)^3}$ なので，

   $$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\ &= \phi''(0) - \big(\phi'(0)\big)^2 \\ &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}\end{aligned}$$

χ^2 分布の積率母関数： $\phi(\theta) = \left(\dfrac{1}{1 - 2\theta}\right)^{\frac{n}{2}}$

$\chi^2(n)$ 分布

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}\exp(-\frac{x}{2}), \quad 0 < x$$

積率母関数

$$\begin{aligned}&\phi(\theta) \\ &= \int_{-\infty}^{\infty} e^{\theta x} f(x)\mathrm{d}x \\ &= \int_{0}^{\infty} e^{\theta x} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}\exp(-\frac{x}{2})\mathrm{d}x \\ &= \int_{0}^{\infty} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}\exp\left(-\frac{1}{2}(1-2\theta)x\right)\mathrm{d}x \\ &= \int_{0}^{\infty} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \left(\frac{y}{1-2\theta}\right)^{\frac{n}{2}-1}\exp(-\frac{1}{2}y)\frac{1}{1-2\theta}\mathrm{d}y \\ &= \left(\frac{1}{1-2\theta}\right)^{\frac{n}{2}}\int_{0}^{\infty} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1}\exp(-\frac{1}{2}y)\mathrm{d}y \\ &= \left(\frac{1}{1-2\theta}\right)^{\frac{n}{2}}\end{aligned}$$

$y = (1-2\theta)x$ として置換積分 ($\dfrac{\mathrm{d}x}{\mathrm{d}y} = (1-2\theta)^{-1}$)

積分のところは，自由度 n の $\chi^2(n)$ 分布に注意

1. 平均：

   $\mathrm{E}(X) = \phi'(0)$

   $\phi(\theta) = (1-2\theta)^{-\frac{n}{2}}$,

   $\phi'(\theta) = (-\frac{n}{2})(-2)(1-2\theta)^{-\frac{n}{2}-1}$ なので，

   $$\begin{aligned}\mathrm{E}(X) &= \phi'(0) \\ &= n\end{aligned}$$

2. 分散：

   $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

   $\mathrm{E}(X^2) = \phi''(0)$

   $\phi''(\theta) = (-\frac{n}{2})(-\frac{n}{2}-1)(-2)^2(1-2\theta)^{-\frac{n}{2}-1}$ なので，

   $$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\ &= \phi''(0) - \big(\phi'(0)\big)^2 \\ &= n(n+2) - n^2 \\ &= 2n\end{aligned}$$

例題： ベルヌイ分布の和の分布： $X_1, X_2, \cdots, X_n$ はそれぞれ独立に同一のベルヌイ分布に従うものとする。このとき，$Y = X_1 + X_2 + \cdots + X_n$ は二項分布 $b(y; n, p)$ に従う。

証明：
$P(X_i=1)=p$ のとき，X_i の積率母関数 $\phi_i(\theta)$ は,

$$\phi_i(\theta)=pe^{\theta}+1-p$$

Y の積率母関数 $\phi_y(\theta)$ は,

$$\begin{aligned}\phi_y(\theta)&=\mathrm{E}(e^{\theta Y})\\&=\mathrm{E}(e^{\theta(X_1+X_2+\cdots+X_n)})\\&=\mathrm{E}(e^{\theta X_1})\mathrm{E}(e^{\theta X_2})\cdots\mathrm{E}(e^{\theta X_n})\\&=\phi_1(\theta)\phi_2(\theta)\cdots\phi_n(\theta)\\&=\Big(\phi(\theta)\Big)^n\\&=(pe^{\theta}+1-p)^n\end{aligned}$$

これは，二項分布 $b(y;n,p)$ の積率母関数に一致する。
注)
3 つ目の等式は，X_1, X_2, $\cdots$, X_n はそれぞれ独立であるため。
5 つ目の等式は，$X_1, X_2, \cdots, X_n$ は同一の分布に従うため。

例題：正規分布の和の分布： X, Y はたがいに独立な確率変数で，$X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$ とする。このとき，定数 a, b について，$aX+bY\sim N(a\mu_1+b\mu_2, a^2\sigma_1^2+b^2\sigma_2^2)$ となる。
証明：
X, Y の積率母関数 $\phi_x(\theta)$, $\phi_y(\theta)$ は,

$$\begin{aligned}\phi_x(\theta)&=\exp\Big(\mu_1\theta+\frac{1}{2}\sigma_1^2\theta^2\Big)\\\phi_y(\theta)&=\exp\Big(\mu_2\theta+\frac{1}{2}\sigma_2^2\theta^2\Big)\end{aligned}$$

$W=aX+bY$ の積率母関数 $\phi_w(\theta)$ は,

$$\begin{aligned}\phi_w(\theta)&=\mathrm{E}(e^{\theta W})\\&=\mathrm{E}(e^{\theta(aX+bY)})\\&=\mathrm{E}(e^{a\theta X})\mathrm{E}(e^{b\theta Y})\\&=\phi_x(a\theta)\phi_y(b\theta)\\&=\exp\Big(\mu_1(a\theta)+\frac{1}{2}\sigma_1^2(a\theta)^2\Big)\\&\quad\times\exp\Big(\mu_2(b\theta)+\frac{1}{2}\sigma_2^2(b\theta)^2\Big)\\&=\exp\Big((a\mu_1+b\mu_2)\theta+\frac{1}{2}(a^2\sigma_1^2+b^2\sigma_2^2)\theta^2\Big)\end{aligned}$$

これは，平均 $a\mu_1+b\mu_2$，分散 $a^2\sigma^2+b^2\sigma_2^2$ の正規分布の積率母関数に一致する。
よって，$aX+bY\sim N(a\mu_1+b\mu_2, a^2\sigma^2+b^2\sigma_2^2)$ となる。

例題：χ^2 分布の和の分布 X, Y は互いに独立で，$X\sim\chi^2(n)$, $Y\sim\chi^2(m)$ とするとき，$Z=X+Y\sim\chi^2(n+m)$ となる。
証明：
X, Y の積率母関数 $\phi_x(\theta), \phi_y(\theta)$ とする。このとき，$\phi_x(\theta)$, $\phi_y(\theta)$ はそれぞれ,

$$\begin{aligned}\phi_x(\theta)&=\Big(\frac{1}{1-2\theta}\Big)^{\frac{n}{2}},\\\phi_y(\theta)&=\Big(\frac{1}{1-2\theta}\Big)^{\frac{m}{2}},\end{aligned}$$

となる。$Z=X+Y$ の積率母関数 $\phi_z(t)$ は，X, Y は互いに独立な確率変数なので,

$$\begin{aligned}\phi_z(\theta)&\equiv\mathrm{E}(e^{\theta Z})\\&=\mathrm{E}(e^{\theta(X+Y)})\\&=\mathrm{E}(e^{\theta X})\mathrm{E}(e^{\theta Y})\\&=\phi_x(\theta)\phi_y(\theta)\\&=\Big(\frac{1}{1-2\theta}\Big)^{\frac{n}{2}}\Big(\frac{1}{1-2\theta}\Big)^{\frac{m}{2}}\\&=\Big(\frac{1}{1-2\theta}\Big)^{\frac{n+m}{2}}\end{aligned}$$

これは，自由度 $n+m$ の χ^2 分布の積率母関数に等しい。したがって，$Z\sim\chi^2(n+m)$ となる。
ただし，3 つ目の等号が成り立つ理由は，X と Y は独立な確率変数であるためである。

5.2 積率母関数 (多変数)

確率変数 X, Y について,
積率母関数： $\phi(\theta_1,\theta_2)=\mathrm{E}(e^{\theta_1X+\theta_2Y})$
性質：

1. 多変数の積率

$$\frac{\partial^{j+k}\phi(0,0)}{\partial\theta_1^j\partial\theta_2^k}=\mathrm{E}(X^jY^k)$$

2. (X_1,Y_1) の積率母関数と (X_2,Y_2) の積率母関数が一致すれば，(X_1,Y_1) の分布関数と (X_2,Y_2) の分布関数も一致する。

3. (X,Y) の積率母関数 $\phi(\theta_1,\theta_2)$,

   X の積率母関数 $\phi_1(\theta_1)$,

Y の積率母関数 $\phi_2(\theta_2)$ について，

$$\phi_1(\theta_1) = \phi(\theta_1, 0), \quad \phi_2(\theta_2) = \phi(0, \theta_2)$$

4. (X, Y) の積率母関数 $\phi(\theta_1, \theta_2)$,
   X の積率母関数 $\phi_1(\theta_1)$,
   Y の積率母関数 $\phi_2(\theta_2)$ について，
   X と Y が独立であるための条件は,

$$\phi(\theta_1, \theta_2) = \phi_1(\theta_1)\phi_2(\theta_2)$$

確率変数 X_1, X_2, $\cdots$, X_n について，
積率母関数：$\phi(\theta_1, \theta_2, \cdots, \theta_n) = \mathrm{E}(e^{\theta_1 X_1 + \theta_2 X_2 + \cdots + \theta_n X_n})$

2 変数正規分布の積率母関数： (X, Y) の確率密度関数

$$\begin{aligned}
&f(x, y) \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\
&\times \exp\Bigg(-\frac{1}{2(1-\rho^2)}\Big(\frac{(x-\mu_1)^2}{\sigma_1^2} \\
&\qquad -2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \\
&\qquad +\frac{(y-\mu_2)^2}{\sigma_2^2}\Big)\Bigg) \\
&= \frac{1}{2\pi}\begin{vmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}^{-1} \\
&\times \exp\left(-\frac{1}{2}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}' \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix} \right) \\
&= \frac{1}{2\pi}|\Sigma|^{-1}\exp(-\frac{1}{2}\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x})
\end{aligned}$$

積率母関数

$$\begin{aligned}
&\phi(\theta_1, \theta_2) \\
&= \mathrm{E}(e^{\theta_1 X + \theta_2 Y}) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\theta_1 x + \theta_2 y} f(x, y)\mathrm{d}x\mathrm{d}y \\
&= \iint e^{\theta_1(x-\mu_1)+\theta_2(y-\mu_2)+\theta_1\mu_1+\theta_2\mu_2} f(x, y)\mathrm{d}x\mathrm{d}y \\
&= \iint \exp\Big(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}'\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix} + \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}'\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}\Big) \\
&\qquad\qquad \times f(x, y)\mathrm{d}x\mathrm{d}y \\
&= \int \exp(\theta'\boldsymbol{x} + \theta'\mu) f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \int \frac{1}{2\pi}|\Sigma|^{-1}\exp(-\frac{1}{2}\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x} + \theta'\boldsymbol{x} + \theta'\mu)\mathrm{d}\boldsymbol{x} \\
&= \int \frac{1}{2\pi}|\Sigma|^{-1}\exp(-\frac{1}{2}(\boldsymbol{x}-\theta)'\Sigma^{-1}(\boldsymbol{x}-\theta) \\
&\qquad\qquad +\theta'\mu + \frac{1}{2}\theta'\Sigma\theta)\mathrm{d}\boldsymbol{x} \\
&= \exp(\theta'\mu + \frac{1}{2}\theta'\Sigma\theta) \\
&\times \int \frac{1}{2\pi}|\Sigma|^{-1}\exp(-\frac{1}{2}(\boldsymbol{x}-\theta)'\Sigma^{-1}(\boldsymbol{x}-\theta))\mathrm{d}\boldsymbol{x} \\
&= \exp(\theta'\mu + \frac{1}{2}\theta'\Sigma\theta) \\
&= \exp\Big(\mu_1\theta_1 + \mu_2\theta_2 + \frac{1}{2}(\sigma_1^2\theta_1^2 + 2\sigma_1\sigma_2\rho\theta_1\theta_2 + \sigma_2^2\theta_2^2)\Big)
\end{aligned}$$

ただし，

$$f(\boldsymbol{x}) = f(x, y),$$

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

とする。

1. $\mathrm{E}(X) = \left.\dfrac{\partial\phi(\theta_1, \theta_2)}{\partial\theta_1}\right|_{(\theta_1,\theta_2)=(0,0)}$

2. $\mathrm{E}(Y) = \left.\dfrac{\partial\phi(\theta_1, \theta_2)}{\partial\theta_2}\right|_{(\theta_1,\theta_2)=(0,0)}$

3. $\mathrm{E}(X^2) = \left.\dfrac{\partial^2\phi(\theta_1, \theta_2)}{\partial\theta_1^2}\right|_{(\theta_1,\theta_2)=(0,0)}$

   $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$

4. $\mathrm{E}(Y^2) = \left.\dfrac{\partial^2\phi(\theta_1, \theta_2)}{\partial\theta_2^2}\right|_{(\theta_1,\theta_2)=(0,0)}$

   $\mathrm{V}(Y) = \mathrm{E}(Y^2) - \big(\mathrm{E}(Y)\big)^2$

5. $\mathrm{E}(XY) = \left.\dfrac{\partial^2\phi(\theta_1, \theta_2)}{\partial\theta_1\partial\theta_2}\right|_{(\theta_1,\theta_2)=(0,0)}$

   $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$

練習問題と解答 (1 章〜 5 章)

$\boxed{1}$ 連続型確率変数 X の密度関数が

$$f(x) = \begin{cases} a-x, & 0<x<a \text{ のとき} \\ 0, & \text{その他} \end{cases}$$

であるとき，次の問に答えよ。

(1) a を求めよ。

(2) X の平均と分散を求めよ。

(3) $Y = X^2$ とするとき，Y の密度関数を求めよ。

[解答]

(1) 密度関数の性質 $\int f(x)\mathrm{d}x = 1$ から，

$$\begin{aligned}\int f(x)\mathrm{d}x &= \int_0^a (a-x)\mathrm{d}x \\ &= \left[ax - \frac{1}{2}x^2\right]_0^a \\ &= \frac{1}{2}a^2 \\ &= 1\end{aligned}$$

により，$a = \sqrt{2}$ を得る。($a > 0$ なので)

(2) 平均，分散の定義は，$\mathrm{E}(X) = \int xf(x)\mathrm{d}x$, $\mathrm{V}(X) = \int (x-\mu)^2 f(x)\mathrm{d}x$ (ただし，$\mu = \mathrm{E}(X)$ とする) である。よって，

$$\begin{aligned}\mathrm{E}(X) &= \int xf(x)\mathrm{d}x \\ &= \int_0^a x(a-x)\mathrm{d}x \\ &= \left[\frac{1}{2}ax^2 - \frac{1}{3}x^3\right]_0^a \\ &= \frac{1}{6}a^3 \\ &= \frac{\sqrt{2}}{3} \quad \longleftarrow \quad a = \sqrt{2} \text{ を代入する}\end{aligned}$$

$$\begin{aligned}\mathrm{V}(X) &= \int (x-\mu)^2 f(x)\mathrm{d}x \\ &= \int x^2 f(x)\mathrm{d}x - \mu^2 \\ &= \int_0^a x^2(a-x)\mathrm{d}x - \mu^2 \\ &= \left[\frac{1}{3}ax^3 - \frac{1}{4}x^4\right]_0^a - \mu^2 \\ &= \frac{1}{12}a^4 - \mu^2 \\ &= \frac{1}{3} - \left(\frac{\sqrt{2}}{3}\right)^2 \\ &= \frac{1}{9}\end{aligned}$$

(3) X の密度関数を $f(x)$，分布関数を $F(x)$ とする。また，Y の密度関数を $g(y)$ とし，分布関数を $G(y)$ とする。$Y = X^2$ なので，

$$\begin{aligned}&G(y) \\ &= P(Y<y) \\ &= P(X^2<y) \\ &= P(-\sqrt{y}<X<\sqrt{y}) \\ &= F(\sqrt{y}) - F(-\sqrt{y}) \\ &= F(\sqrt{y}) \quad \longleftarrow \quad F(-\sqrt{y}) = 0\end{aligned}$$

を得る。さらに，密度関数と分布関数の関係から，

$$\begin{aligned}g(y) &= \frac{\mathrm{d}G(y)}{\mathrm{d}y} \\ &= \frac{\mathrm{d}F(\sqrt{y})}{\mathrm{d}y} \\ &= \frac{\mathrm{d}F(x)}{\mathrm{d}x}\frac{\mathrm{d}\sqrt{y}}{\mathrm{d}y} \quad \longleftarrow \quad x = \sqrt{y} \\ &= F'(x)\frac{1}{2\sqrt{y}} \\ &= f(x)\frac{1}{2\sqrt{y}} \\ &= f(\sqrt{y})\frac{1}{2\sqrt{y}} \\ &= (\sqrt{2}-\sqrt{y})\frac{1}{2\sqrt{y}} \qquad 0<y<2 \text{ のとき}\end{aligned}$$

y の範囲は，

$0<x<\sqrt{2} \Longrightarrow 0<x^2<2 \Longrightarrow 0<y<2$

となる。

$\boxed{2}$ 連続型確率変数 X の密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$$

であるとき，次の問に答えよ。

(1) X の平均と分散を求めよ。

(2) $Y = X^2$ とするとき，Y の平均と分散を求めよ。

(3) $Z = e^X$ とするとき，Z の平均と分散を求めよ。

[解答]

(1) 平均，分散の定義は，$\mathrm{E}(X) = \displaystyle\int xf(x)\mathrm{d}x$, $\mathrm{V}(X) = \displaystyle\int (x-\mu)^2 f(x)\mathrm{d}x$ (ただし，$\mu = \mathrm{E}(X)$ とする) である。よって，

$$\begin{aligned}
\mathrm{E}(X) &= \int xf(x)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\mathrm{d}x \\
&= -\frac{1}{\sqrt{2\pi}}\left[e^{-\frac{1}{2}x^2}\right]_{-\infty}^{\infty} \\
&= 0
\end{aligned}$$

3 つ目の等式は，$\dfrac{\mathrm{d}e^{-\frac{1}{2}x^2}}{\mathrm{d}x} = -xe^{-\frac{1}{2}x^2}$ を利用する。

$$\begin{aligned}
\mathrm{V}(X) &= \int (x-\mu)^2 f(x)\mathrm{d}x \\
&= \int x^2 f(x)\mathrm{d}x - \mu^2 \\
&= \int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\mathrm{d}x - \mu^2 \\
&= \left[-x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\right]_{-\infty}^{\infty} \\
&\quad + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\mathrm{d}x - \mu^2 \\
&= 1
\end{aligned}$$

4 つ目の等式では，部分積分を利用

$$\begin{aligned}
&\int_a^b h'(x)g(x)\mathrm{d}x \\
&= [h(x)g(x)]_a^b - \int_a^b h(x)g'(x)\mathrm{d}x
\end{aligned}$$

$g(x) = x$, $h'(x) = x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ とおく。

また，4 つ目の等式の第 1 項では，

$$\lim_{x\to\pm\infty} x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} = 0$$

を利用する。

4 つ目の等式の第 2 項では，密度関数の積分が 1 になることを利用。

(2) $Y = X^2$ とするとき，Y の平均と分散を求める。

$$\begin{aligned}
\mathrm{E}(Y) &= \mathrm{E}(X^2) \\
&= \mathrm{V}(X) - \mu_x^2 \\
&= 1
\end{aligned}$$

(1) より，$\mathrm{V}(X) = 1$, $\mu_x = \mathrm{E}(X) = 0$ に注意。

$$\begin{aligned}
&\mathrm{V}(Y) \\
&= \mathrm{E}(Y-\mu_y)^2 \quad \longleftarrow \quad \mu_y = \mathrm{E}(Y) = 1 \\
&= \mathrm{E}(Y^2) - \mu_y^2 \\
&= \mathrm{E}(X^4) - \mu_y^2 \\
&= \int_{-\infty}^{\infty} x^4\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\mathrm{d}x - \mu_y^2 \\
&= \int_{-\infty}^{\infty} x^3\cdot x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\mathrm{d}x - \mu_y^2 \\
&= \left[-x^3\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\right]_{-\infty}^{\infty} \\
&+ 3\int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\mathrm{d}x - \mu_y^2 \\
&= 3\mathrm{E}(X^2) - \mu_y^2 \quad \longleftarrow \quad \mathrm{E}(X^2) = 1,\ \mu_y = 1 \\
&= 2
\end{aligned}$$

6 つ目の等式では，部分積分を利用

$$\begin{aligned}
&\int_a^b h'(x)g(x)\mathrm{d}x \\
&= [h(x)g(x)]_a^b - \int_a^b h(x)g'(x)\mathrm{d}x
\end{aligned}$$

$g(x) = x^3$, $h'(x) = x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ とおく。

また，6 つ目の等式の第 1 項では，

$$\lim_{x\to\pm\infty} x^3\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} = 0$$

を利用する。

(3) $Z = e^X$ とするとき，Z の平均と分散を求める。

$$\mathrm{E}(Z) = \mathrm{E}(e^X)$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} e^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2-2x)} \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2+\frac{1}{2}} \mathrm{d}x \\
&= e^{\frac{1}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2} \mathrm{d}x \\
&= e^{\frac{1}{2}}
\end{aligned}
$$

6 つ目の等式は，$\dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2}$ が，平均 1，分散 1 の正規分布となり，その積分値は 1 となることによる。

$$
\begin{aligned}
\mathrm{V}(Z) &= \mathrm{E}(Z-\mu_z)^2 \quad \longleftarrow \quad \mu_z = \mathrm{E}(Z) = e^{\frac{1}{2}} \\
&= \mathrm{E}(Z^2) - \mu_z^2 \\
&= \mathrm{E}(e^{2X}) - \mu_z^2 \\
&= \int_{-\infty}^{\infty} e^{2x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \mathrm{d}x - \mu_z^2 \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2-4x)} \mathrm{d}x - \mu_z^2 \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-2)^2+2} \mathrm{d}x - \mu_z^2 \\
&= e^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-2)^2} \mathrm{d}x - \mu_z^2 \\
&= e^2 - e
\end{aligned}
$$

8 つ目の等式は，$\dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-2)^2}$ が，平均 2，分散 1 の正規分布となり，その積分値は 1 となることによる。

$\boxed{3}$ 連続型確率変数 X の密度関数が

$$
f(x) = \begin{cases} \dfrac{1}{\lambda} e^{-\frac{x}{\lambda}}, & 0 < x \text{ のとき} \\ 0, & \text{その他} \end{cases}
$$

であるとき，次の問に答えよ。

(1) X の平均と分散を求めよ。

(2) X の積率母関数を求めよ。

(3) $X_1, X_2, \cdots, X_n$ を互いに独立で上に示された分布に従うものとする。$\lambda = 2$ のとき，$Y = X_1 + X_2 + \cdots + X_n$ の密度関数は，自由度 $2n$ のカイ二乗分布となることを示せ。ただし，自由度 n のカイ二乗分布とは $\boxed{5}$ の確率変数 X の密度関数である。

[解答]

(1) X の平均と分散を求める。

$$
\begin{aligned}
\mathrm{E}(X) &= \int x f(x) \mathrm{d}x \\
&= \int_0^{\infty} x \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \mathrm{d}x \\
&= \left[-x e^{-\frac{x}{\lambda}}\right]_0^{\infty} \\
&\quad + \int_0^{\infty} e^{-\frac{x}{\lambda}} \mathrm{d}x \\
&= \left[-\lambda e^{-\frac{x}{\lambda}}\right]_0^{\infty} \\
&= \lambda
\end{aligned}
$$

3 つ目の等式では，部分積分を利用

$$
\begin{aligned}
&\int_a^b h'(x) g(x) \mathrm{d}x \\
&= [h(x)g(x)]_a^b - \int_a^b h(x) g'(x) \mathrm{d}x
\end{aligned}
$$

$g(x) = x$, $h'(x) = \dfrac{1}{\lambda} e^{-\frac{x}{\lambda}}$ とおく。

また，

$$
\lim_{x \to \infty} x e^{-\frac{x}{\lambda}} = 0
$$

$$
\lim_{x \to \infty} e^{-\frac{x}{\lambda}} = 0
$$

を利用

$$
\begin{aligned}
&\mathrm{V}(X) \\
&= \int (x-\mu)^2 f(x) \mathrm{d}x \\
&= \int x^2 f(x) \mathrm{d}x - \mu^2 \quad \longleftarrow \quad \mu = \mathrm{E}(X) = \lambda \\
&= \int_0^{\infty} x^2 \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \mathrm{d}x - \mu^2 \\
&= \left[-x^2 e^{-\frac{x}{\lambda}}\right]_0^{\infty} + 2 \int_0^{\infty} x e^{-\frac{x}{\lambda}} \mathrm{d}x - \mu^2 \\
&= \left[-x^2 e^{-\frac{x}{\lambda}}\right]_0^{\infty} + 2\lambda \int_0^{\infty} x \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \mathrm{d}x - \mu^2 \\
&= 2\lambda \mathrm{E}(X) - \mu^2 \quad \longleftarrow \quad \mu = \mathrm{E}(X) = \lambda \\
&= 2\lambda^2 - \lambda^2 \\
&= \lambda^2
\end{aligned}
$$

3 つ目の等式では，部分積分を利用

$$\int_a^b h'(x)g(x)\mathrm{d}x$$
$$= [h(x)g(x)]_a^b - \int_a^b h(x)g'(x)\mathrm{d}x$$

$g(x) = x^2$, $h'(x) = \dfrac{1}{\lambda}e^{-\frac{x}{\lambda}}$ とおく。

6 つ目の等式では，

$$\lim_{x\to\infty} x^2 e^{-\frac{x}{\lambda}} = 0$$

$$\mu = \mathrm{E}(X) = \int_0^\infty x e^{-\frac{x}{\lambda}}\mathrm{d}x$$

を利用。

(2) X の積率母関数を求める。

$$\begin{aligned}\phi(\theta) &= \mathrm{E}(e^{\theta X})\\ &= \int e^{\theta x} f(x)\mathrm{d}x\\ &= \int_0^\infty e^{\theta x}\frac{1}{\lambda}e^{-\frac{x}{\lambda}}\mathrm{d}x\\ &= \int_0^\infty \frac{1}{\lambda}e^{-(\frac{1}{\lambda}-\theta)x}\mathrm{d}x\\ &= \frac{\frac{1}{\lambda}}{\frac{1}{\lambda}-\theta}\int_0^\infty (\frac{1}{\lambda}-\theta)e^{-(\frac{1}{\lambda}-\theta)x}\mathrm{d}x\\ &= \frac{1}{1-\lambda\theta}\end{aligned}$$

最後の等式では，$(\dfrac{1}{\lambda}-\theta)e^{-(\frac{1}{\lambda}-\theta)x}$ は密度関数であるので，その積分値は 1 であることによる。$f(x)$ の λ を $\dfrac{1}{\lambda}-\theta$ で置き換えたものとなっている。

(3) Y の積率母関数と自由度 $2n$ のカイ二乗分布の積率母関数が一致することを示す。

X_1, X_2, $\cdots$, X_n を互いに独立で上に示された分布に従うので，X_i の積率母関数 $\phi_i(\theta)$ は，(2) より，$\lambda = 2$ のとき，

$$\phi_i(\theta) = \frac{1}{1-2\theta} = \phi(\theta)$$

となる。

$\lambda = 2$ のとき，$Y = X_1 + X_2 + \cdots + X_n$ の積率母関数 $\phi_y(\theta)$ は，

$$\begin{aligned}\phi_y(\theta) &= \mathrm{E}(e^{\theta Y})\\ &= \mathrm{E}(e^{\theta(X_1+X_2+\cdots+X_n)})\\ &= \mathrm{E}(e^{\theta X_1})\mathrm{E}(e^{\theta X_2})\cdots\mathrm{E}(e^{\theta X_n})\\ &= \phi_1(\theta)\phi_2(\theta)\cdots\phi_n(\theta)\\ &= \Big(\phi(\theta)\Big)^n\\ &= \Big(\frac{1}{1-2\theta}\Big)^n\\ &= \Big(\frac{1}{1-2\theta}\Big)^{\frac{2n}{2}}\end{aligned}$$

したがって，Y の積率母関数は，

$$\phi_y(\theta) = \Big(\frac{1}{1-2\theta}\Big)^{\frac{2n}{2}}$$

となる。

一方，自由度 m のカイ二乗分布は，

$$f(x) = \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})}x^{\frac{m}{2}-1}e^{-\frac{x}{2}} \quad x > 0 \text{ のとき}$$

なので，その積率母関数 $\phi_{\chi^2}(\theta)$ は，

$$\begin{aligned}&\phi_{\chi^2}(\theta)\\ &= \mathrm{E}(e^{\theta X})\\ &= \int_0^\infty e^{\theta x}\frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})}x^{\frac{m}{2}-1}e^{-\frac{x}{2}}\mathrm{d}x\\ &= \int_0^\infty \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})}x^{\frac{m}{2}-1}e^{-\frac{1}{2}(1-2\theta)x}\mathrm{d}x\\ &= \int_0^\infty \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})}\left(\frac{y}{1-2\theta}\right)^{\frac{m}{2}-1}e^{-\frac{1}{2}y}\frac{1}{1-2\theta}\mathrm{d}x\\ &= \left(\frac{1}{1-2\theta}\right)^{\frac{m}{2}-1}\frac{1}{1-2\theta}\\ &\times\int_0^\infty \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})}y^{\frac{m}{2}-1}e^{-\frac{1}{2}y}\mathrm{d}x\\ &= \left(\frac{1}{1-2\theta}\right)^{\frac{m}{2}}\end{aligned}$$

4 つ目の等式で，$y = (1-2\theta)x$ として，置換積分を利用。

$\dfrac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})}y^{\frac{m}{2}-1}e^{-\frac{1}{2}y}$ は，自由度 m の χ^2 分布となっているので，その積分値は 1 となる。

$\phi_y(\theta)$ は，$\phi_{\chi^2}(\theta)$ で，$m = 2n$ に対応する。

すなわち，$\phi_y(\theta)$ は自由度 $2n$ の χ^2 分布の積率母関数となっている。

したがって，$Y \sim \chi^2(2n)$ となる。

$\boxed{4}$ 連続型確率変数 X の密度関数が

$$f(x) = \begin{cases} 1, & 0 < x < 1 \text{ のとき} \\ 0, & \text{その他} \end{cases}$$

であるとき，次の問に答えよ。

(1) X の平均と分散を求めよ。

(2) $Y = -2\log X$ とするとき，Y の積率母関数を求めよ。ただし，log は自然対数とする。($y = -2\log x$ は $x = e^{-\frac{1}{2}y}$ を意味する)

(3) Y_1 と Y_2 を (2) で求められた密度関数に従う確率変数であるとする。しかも，Y_1 と Y_2 は独立であるとする。$Z = Y_1 + Y_2$ としたとき，Z の密度関数を求めよ。

[解答]

(1) X の平均と分散を求める。

$$\begin{aligned} \mathrm{E}(X) &= \int x f(x)\mathrm{d}x \\ &= \int_0^1 x\mathrm{d}x \\ &= \Big[\frac{1}{2}x^2\Big]_0^1 \\ &= \frac{1}{2} \end{aligned}$$

$$\begin{aligned} &\mathrm{V}(X) \\ &= \int (x-\mu)^2 f(x)\mathrm{d}x \\ &= \int x^2 f(x)\mathrm{d}x - \mu^2 \qquad \longleftarrow \qquad \mu = \mathrm{E}(X) = \frac{1}{2} \\ &= \int_0^1 x^2\mathrm{d}x - \mu^2 \\ &= \Big[\frac{1}{3}x^3\Big]_0^1 - \mu^2 \\ &= \frac{1}{3} - \Big(\frac{1}{2}\Big)^2 \\ &= \frac{1}{12} \end{aligned}$$

(2) $Y = -2\log X$ とするとき，Y の積率母関数 $\phi_y(\theta)$ を求める。

$$\begin{aligned} \phi_y(\theta) &= \mathrm{E}(e^{\theta Y}) \\ &= \mathrm{E}(e^{-2\theta\log X}) \\ &= \mathrm{E}(X^{-2\theta}) \\ &= \int x^{-2\theta} f(x)\mathrm{d}x \\ &= \int_0^1 x^{-2\theta}\mathrm{d}x \\ &= \Big[\frac{1}{1-2\theta}x^{1-2\theta}\Big]_0^1 \\ &= \frac{1}{1-2\theta} \end{aligned}$$

(3) Y_1 と Y_2 を (2) で求められた密度関数に従う確率変数であるとする。しかも，Y_1 と Y_2 は独立であるとする。$Z = Y_1 + Y_2$ としたとき，Z の密度関数を求める。

Z の積率母関数 $\phi_z(\theta)$ を求める。

$$\begin{aligned} \phi_z(\theta) &= \mathrm{E}(e^{\theta Z}) \\ &= \mathrm{E}(e^{\theta(Y_1+Y_2)}) \\ &= \mathrm{E}(e^{\theta Y_1})\mathrm{E}(e^{\theta Y_2}) \\ &= \Big(\phi_y(\theta)\Big)^2 \\ &= \Big(\frac{1}{1-2\theta}\Big)^2 \\ &= \Big(\frac{1}{1-2\theta}\Big)^{\frac{4}{2}} \end{aligned}$$

これは，自由度 4 のカイ自乗分布の積率母関数に一致する。

よって，$Z \sim \chi^2(4)$ となる。

ただし，自由度 n のカイ自乗分布の密度関数は,

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \text{ のとき} \\ 0, & \text{その他} \end{cases}$$

として表され，その積率母関数 $\phi(\theta)$ は,

$$\phi(\theta) = \Big(\frac{1}{1-2\theta}\Big)^{\frac{n}{2}}$$

となることに注意。

$\boxed{5}$ 連続型確率変数 X の密度関数が

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \text{ のとき} \\ 0, & \text{その他} \end{cases}$$

であるとき，次の問に答えよ。ただし，$\Gamma(a)$ はガンマ関数であり，

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} \mathrm{d}x$$

と定義される。

(1) X の平均と分散を求めよ。

(2) X の積率母関数を求めよ。

[**解答**]

(1) X の平均と分散を求める。

平均について：

$$\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^\infty x f(x) \mathrm{d}x \\
&= \int_0^\infty x \frac{1}{\Gamma(\frac{n}{2})} 2^{-\frac{n}{2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= \frac{2^{-\frac{n}{2}}}{2^{-\frac{n+2}{2}}} \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n}{2})} \\
&\quad \times \int_0^\infty \frac{1}{\Gamma(\frac{n+2}{2})} 2^{-\frac{n+2}{2}} x^{\frac{n+2}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= 2\frac{n}{2} \int_0^\infty \frac{1}{\Gamma(\frac{n'}{2})} 2^{-\frac{n'}{2}} x^{\frac{n'}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= n
\end{aligned}$$

$\Gamma(s) = \displaystyle\int_0^\infty u^{s-1} e^{-u} \mathrm{d}u \Longrightarrow$ ガンマ関数 $\Gamma(s+1) = s\Gamma(s), \Gamma(1) = 1, \Gamma(\dfrac{1}{2}) = \sqrt{\pi}$ に注意

また，$n' = n + 2$ を使い，確率密度関数の性質から，

$$\int_{-\infty}^\infty f(x) \mathrm{d}x = \int_0^\infty \frac{1}{\Gamma(\frac{n}{2})} 2^{-\frac{n}{2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x = 1$$

に注意。

分散について：

$\mathrm{V}(X) = \mathrm{E}(X^2) - \mu^2$ により，$\mathrm{E}(X^2)$ を求める。

$$\begin{aligned}
&\mathrm{E}(X^2) \\
&= \int_{-\infty}^\infty x^2 f(x) \mathrm{d}x \\
&= \int_0^\infty x^2 \frac{1}{\Gamma(\frac{n}{2})} 2^{-\frac{n}{2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= \int_0^\infty \frac{1}{\Gamma(\frac{n}{2})} 2^{-\frac{n}{2}} x^{\frac{n+4}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= \frac{2^{-\frac{n}{2}}}{2^{-\frac{n+4}{2}}} \frac{\Gamma(\frac{n+4}{2})}{\Gamma(\frac{n}{2})} \\
&\quad \times \int_0^\infty \frac{1}{\Gamma(\frac{n+4}{2})} 2^{-\frac{n+4}{2}} x^{\frac{n+4}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= 4(\frac{n+2}{2}\frac{n}{2}) \int_0^\infty \frac{1}{\Gamma(\frac{n'}{2})} 2^{-\frac{n'}{2}} x^{\frac{n'}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x \\
&= n(n+2)
\end{aligned}$$

$n' = n + 4$ を使う。

$\mathrm{V}(X) = n(n+2) - n^2 = 2n$

(2) X の積率母関数を求める。

$$\begin{aligned}
&\phi(\theta) \\
&= \mathrm{E}(e^{\theta X}) \\
&= \int_{-\infty}^\infty e^{\theta x} f(x) \mathrm{d}x \\
&= \int_0^\infty e^{\theta x} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} \exp(-\frac{x}{2}) \mathrm{d}x \\
&= \int_0^\infty \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} \exp\left(-\frac{1}{2}(1-2\theta)x\right) \mathrm{d}x \\
&= \int_0^\infty \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \left(\frac{y}{1-2\theta}\right)^{\frac{n}{2}-1} \exp(-\frac{1}{2}y) \frac{1}{1-2\theta} \mathrm{d}y \\
&= \left(\frac{1}{1-2\theta}\right)^{\frac{n}{2}} \int_0^\infty \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} \exp(-\frac{1}{2}y) \mathrm{d}y \\
&= \left(\frac{1}{1-2\theta}\right)^{\frac{n}{2}}
\end{aligned}$$

$y = (1-2\theta)x$ として置換積分 $(\dfrac{\mathrm{d}x}{\mathrm{d}y} = (1-2\theta)^{-1})$

積分のところは，自由度 n の $\chi^2(n)$ 分布に注意

$\boxed{6}$ 連続型確率変数 X, Y は互いに独立で，$X \sim N(0,1)$，$Y \sim N(0,1)$ とする。$U = \dfrac{X}{Y}$ とするとき，次の問に答えよ。ただし，$X \sim N(0,1)$ のとき，X の密度関数は

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

と書き表される。

(1) U の密度関数を求めよ。

(2) U の 1 次の積率は存在しないということを証明せよ。

[解答]

(1) U の密度関数を求める。

X, Y の密度関数は，それぞれ，

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2), \quad -\infty < x < \infty \\
g(y) &= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2), \quad -\infty < y < \infty
\end{aligned}
$$

となる。

X, Y の結合確率密度関数は，X, Y は互いに独立な確率変数なので，

$$
\begin{aligned}
&h(x, y) \\
&= f(x)g(y) \\
&= \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2) \\
&= \frac{1}{2\pi}\exp\bigl(-\frac{1}{2}(x^2+y^2)\bigr)
\end{aligned}
$$

$u = \dfrac{x}{y}$, $v = y$ として，変数変換を行う。

$x = uv$, $y = v$ なので，

$$
\begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} v & u \\ 0 & 1 \end{pmatrix}
$$

となるので，U, V の結合確率密度関数 $s(u, v)$ は，変数変換により，

$$
\begin{aligned}
&s(u, v) \\
&= h(uv, v)\begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} \\
&= \frac{1}{2\pi}\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)|v|
\end{aligned}
$$

U の周辺確率密度関数は，

$$
\begin{aligned}
&p(u) \\
&= \int s(u, v)\mathrm{d}v \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}|v|\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)\mathrm{d}v \\
&= \frac{1}{\pi}\int_{0}^{\infty}v\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)\mathrm{d}v \\
&= \frac{1}{\pi}\left[-\frac{1}{1+u^2}\exp\bigl(-\frac{1}{2}v^2(1+u^2)\bigr)\right]_{v=0}^{\infty} \\
&= \frac{1}{\pi(1+u^2)}
\end{aligned}
$$

これは，コーシー分布の密度関数である。

(2) U の 1 次の積率 (すなわち，平均) は存在しないということを証明する。

$$
\begin{aligned}
&\mathrm{E}(U) \\
&= \int uf(u)\mathrm{d}u \\
&= \int_{-\infty}^{\infty}u\frac{1}{\pi(1+u^2)}\mathrm{d}u \\
&= \int_{1}^{\infty}\frac{1}{2\pi}\frac{1}{x}\mathrm{d}x \quad \longleftarrow \quad x = 1+u^2 \text{ で置換積分} \\
&= \left[\frac{1}{2\pi}\log x\right]_{1}^{\infty} \quad \longleftarrow \quad \frac{\mathrm{d}\log x}{\mathrm{d}x} = \frac{1}{x} \\
&= \infty
\end{aligned}
$$

$-\infty < u < \infty$ のとき，$x = 1+u^2$ の範囲は，$1 < x < \infty$ となる。

$\boxed{7}$ 連続型確率変数 X, Y の同時密度関数が

$$
f(x, y) = \begin{cases} x+y, & 0 < x < 1,\ 0 < y < 1 \text{ のとき} \\ 0, & \text{その他} \end{cases}
$$

であるとき，次の問に答えよ。

(1) XY の期待値を求めよ。

(2) X と Y の相関係数を求めよ。

(3) X の周辺密度関数を求めよ。

[解答]

(1) XY の期待値を求める。

$$
\begin{aligned}
\mathrm{E}(XY) &= \int_0^1\int_0^1 xyf(x, y)\mathrm{d}x\mathrm{d}y \\
&= \int_0^1\int_0^1 xy(x+y)\mathrm{d}x\mathrm{d}y
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^1 \left[\frac{1}{3}yx^3 + \frac{1}{2}y^2x^2\right]_0^1 \mathrm{d}y \\
&= \int_0^1 (\frac{1}{3}y + \frac{1}{2}y^2)\mathrm{d}y \\
&= \left[\frac{1}{6}y^2 + \frac{1}{6}y^3\right]_0^1 \\
&= \frac{1}{3}
\end{aligned}
$$

(2) X と Y の相関係数 $\rho = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)\mathrm{V}(Y)}}$ を求める。

$$
\begin{aligned}
\mathrm{E}(X) &= \int_0^1 \int_0^1 xf(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_0^1 \int_0^1 x(x+y)\mathrm{d}x\mathrm{d}y \\
&= \int_0^1 \left[\frac{1}{3}x^3 + \frac{1}{2}yx^2\right]_0^1 \mathrm{d}y \\
&= \int_0^1 (\frac{1}{3} + \frac{1}{2}y)\mathrm{d}y \\
&= \left[\frac{1}{3}y + \frac{1}{4}y^2\right]_0^1 \\
&= \frac{7}{12}
\end{aligned}
$$

$f(x,y)$ の形は，x と y を入れ替えても同じ形なので，

$$
\mathrm{E}(Y) = \mathrm{E}(X) = \frac{7}{12}
$$

となる。

$$
\begin{aligned}
&\mathrm{V}(X) \\
&= \mathrm{E}\Big((X-\mu)^2\Big) \quad \longleftarrow \quad \mu = \mathrm{E}(X) = \frac{7}{12} \\
&= \mathrm{E}(X^2) - \mu^2 \\
&= \int_0^1 \int_0^1 x^2 f(x,y)\mathrm{d}x\mathrm{d}y - \mu^2 \\
&= \int_0^1 \int_0^1 x^2(x+y)\mathrm{d}x\mathrm{d}y - \mu^2 \\
&= \int_0^1 \left[\frac{1}{4}x^4 + \frac{1}{3}yx^3\right]_0^1 \mathrm{d}y - \mu^2 \\
&= \int_0^1 (\frac{1}{4} + \frac{1}{3}y)\mathrm{d}y - \mu^2 \\
&= \left[\frac{1}{4}y + \frac{1}{6}y^2\right]_0^1 - \mu^2 \\
&= \frac{5}{12} - \left(\frac{7}{12}\right)^2 \\
&= \frac{11}{144}
\end{aligned}
$$

同様に，

$$
\mathrm{V}(Y) = \mathrm{V}(X) = \frac{11}{144}
$$

となる。

$$
\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathrm{E}\Big((X-\mu_x)(Y-\mu_y)\Big) \\
&= \mathrm{E}(XY) - \mu_x\mu_y \\
&= \frac{1}{3} - \frac{7}{12}\frac{7}{12} \\
&= -\frac{1}{144}
\end{aligned}
$$

ただし，

$$
\mu_x = \mathrm{E}(X) = \frac{7}{12}, \quad \mu_y = \mathrm{E}(Y) = \frac{7}{12}
$$

よって，

$$
\begin{aligned}
\rho &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)\mathrm{V}(Y)}} \\
&= \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144}\frac{11}{144}}} = -\frac{1}{11}
\end{aligned}
$$

(3) X の周辺密度関数 $f_x(x)$ を求める。

$$
\begin{aligned}
f_x(x) &= \int f(x,y)\mathrm{d}y \\
&= \int_0^1 (x+y)\mathrm{d}y \\
&= \left[xy + \frac{1}{2}y^2)\right]_{y=0}^1 \\
&= x + \frac{1}{2}
\end{aligned}
$$

$\boxed{8}$ 離散型確率変数 X の密度関数が

$$
f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \cdots,
$$

であるとき，次の問に答えよ。

(1) $\displaystyle\sum_{x=0}^{\infty} f(x) = 1$ となることを証明せよ。

(2) X の積率母関数を求めよ。

(3) 積率母関数をもとにして，X の平均と分散を求めよ。

[解答]

(1) $\displaystyle\sum_{x=0}^{\infty} f(x) = 1$ となることを証明する。

$$\begin{aligned}\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!} \\ &= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} \\ &= e^{-\lambda}e^{\lambda} \\ &= 1\end{aligned}$$

$e^x = \displaystyle\sum_{k=0}^{\infty}\frac{x^k}{k!}$ に注意。

なぜなら，$f(x) = e^x$ としたとき，$f^{(k)}(x) = e^x$ となる。

テーラー展開の公式は，

$$f(x) = \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(x_0)(x-x_0)^k$$

なので，$x_0 = 0$ として，$x = 0$ の回りでテーラー展開すると，

$$\begin{aligned}f(x) &= \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)x^k \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}x^k \\ &= \sum_{k=0}^{\infty}\frac{x^k}{k!}\end{aligned}$$

を得る。

x を λ，k を x で置き換える。

(2) X の積率母関数を求める。

$$\begin{aligned}\phi(\theta) &= \mathrm{E}(e^{\theta X}) \\ &= \sum_{x=0}^{\infty} e^{\theta x} f(x) \\ &= \sum_{x=0}^{\infty} e^{\theta x} e^{-\lambda}\frac{\lambda^x}{x!} \\ &= \sum_{x=0}^{\infty} e^{-\lambda}\frac{(e^{\theta}\lambda)^x}{x!} \\ &= e^{-\lambda}\exp(e^{\theta}\lambda)\sum_{x=0}^{\infty}\exp(-e^{\theta}\lambda)\frac{(e^{\theta}\lambda)^x}{x!} \\ &= e^{-\lambda}\exp(e^{\theta}\lambda)\sum_{x=0}^{\infty} e^{-\lambda'}\frac{\lambda'^x}{x!} \\ &= \exp(-\lambda)\exp(e^{\theta}\lambda) \\ &= \exp\Big(\lambda(e^{\theta}-1)\Big)\end{aligned}$$

ただし，$\lambda' = e^{\theta}\lambda$ に注意

(3) 積率母関数をもとにして，X の平均と分散を求める。

平均について：

$\mathrm{E}(X) = \phi'(0)$

$\phi(\theta) = \exp\Big(\lambda(e^{\theta}-1)\Big)$,

$\phi'(\theta) = \lambda e^{\theta}\exp\Big(\lambda(e^{\theta}-1)\Big)$ なので,

$$\begin{aligned}\mathrm{E}(X) &= \phi'(0) \\ &= \lambda\end{aligned}$$

分散について：

$\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$ なので，$\mathrm{E}(X^2)$ を求める。

$\mathrm{E}(X^2) = \phi''(0)$

$\phi''(\theta) = (1+\lambda e^{\theta})\lambda e^{\theta}\exp\Big(\lambda(e^{\theta}-1)\Big)$ なので,

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\ &= \phi''(0) - \big(\phi'(0)\big)^2 \\ &= (1+\lambda)\lambda - \lambda^2 \\ &= \lambda\end{aligned}$$

6　大数の法則と中心極限定理

6.1　Chebyshev の不等式

$g(x) \geq 0$ について，

$$P\big(g(X) \geq k\big) \leq \frac{\mathrm{E}\big(g(X)\big)}{k}$$

となる。ただし，k は正の定数とする。

証明：

$g(X) \geq k$ のとき $U = 1$，$g(X) < k$ のとき $U = 0$ となる離散型確率変数 U を導入する。

離散型確率変数 U は 0 か 1 の値を取り，その確率関数 $f(u)$ は次のように与えられる。

$$f(u) = P(U = u)$$

ただし，

$$P(U = 1) = P\big(g(X) \geq k\big)$$
$$P(U = 0) = P\big(g(X) < k\big)$$

となる。

このとき，常に，以下の式が成り立つ。

$$g(X) \geq kU$$

よって，両辺に期待値を取ると，

$$\mathrm{E}\big(g(X)\big) \geq k\mathrm{E}(U)$$

となる。$\mathrm{E}(U)$ を求める。

$$\begin{aligned}\mathrm{E}(U) &= \sum_{u=0}^{1} uP(U = u)\\ &= 1 \times P(U = 1) + 0 \times P(U = 0)\\ &= 1 \times P\big(g(X) \geq k\big) + 0 \times P\big(g(X) < k\big)\\ &= P\big(g(X) \geq k\big)\end{aligned}$$

したがって，

$$\mathrm{E}\big(g(X)\big) \geq kP\big(g(X) \geq k\big)$$

から，

$$P\big(g(X) \geq k\big) \leq \frac{\mathrm{E}\big(g(X)\big)}{k}$$

を得る。

代表的な例：　$\mathrm{E}(X) = \mu$, $\mathrm{V}(X) = \sigma^2$, $\lambda > 1$ を任意の定数とする。このとき，

$$P(|X - \mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$$
$$P(|X - \mu| < \lambda\sigma) \geq 1 - \frac{1}{\lambda^2}$$

となる。

証明：

$g(X) = (X - \mu)^2$, $k = \lambda^2\sigma^2$ とすると，

$$P\big(g(X) \geq k\big) \leq \frac{\mathrm{E}\big(g(X)\big)}{k}$$

から，

$$P\big((X - \mu)^2 \geq \lambda^2\sigma^2\big) \leq \frac{\sigma^2}{\lambda^2\sigma^2}$$

$$P\big(|X - \mu| \geq \lambda\sigma\big) \leq \frac{1}{\lambda^2}$$

を得る。

さらに，$\epsilon = \lambda\sigma$ とすると，

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

とも書き表される。

6.2　大数の (弱) 法則 (Convergence in probability)

X_1, X_2, $\cdots$, X_n が互いに独立な確率変数ですべて同じ分布に従うとし，すべての i について $\mathrm{E}(X_i) = \mu$ とする。このとき任意の正数 ϵ について，$n \longrightarrow \infty$ のとき，

$$P(|\overline{X}_n - \mu| > \epsilon) \;\longrightarrow\; 0$$

が成り立つ。ただし，

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

とする。

このとき，$\overline{X}_n$ は μ に確率収束するという。

証明：

$\mathrm{E}(\overline{X}_n) = \mu$，$\mathrm{V}(\overline{X}_n) = \dfrac{\sigma^2}{n}$ なので，

$$P(|\overline{X}_n - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

が成り立つ。したがって，$n \longrightarrow \infty$ のとき，

$$P(|\overline{X}_n - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \ \longrightarrow \ 0$$

となる。すなわち，$\overline{X}_n \longrightarrow \mu$ が得られる。

系：また，$X_1, X_2, \cdots, X_n$ が同じ分布に従わなくても，独立性もなくても，

$$m_n = \mathrm{E}(\sum_{i=1}^{n} X_i),$$

$$V_n = \mathrm{V}(\sum_{i=1}^{n} X_i),$$

が存在し，$n \longrightarrow \infty$ のとき，

$$\frac{V_n}{n^2} \ \longrightarrow \ 0$$

が成り立てば，

$$\frac{\sum_{i=1}^{n} X_i - m_n}{n} \ \longrightarrow \ 0$$

が成り立つ。(証明略)
2 つとも，大数の (弱) 法則と呼ばれる。

6.3 中心極限定理

$X_1, X_2, \cdots, X_n$ が互いに独立な確率変数ですべて同じ分布に従うとし，すべての i について $\mathrm{E}(X_i) = \mu$，$\mathrm{V}(X_i) = \sigma^2$ とする。以上の仮定のもとで，$n \longrightarrow \infty$ のとき，

$$P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} < x\right) \ \longrightarrow \ \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \mathrm{d}u$$

が成り立つ。$\Longrightarrow$ 中心極限定理
証明：
$Y_i = \dfrac{X_i - \mu}{\sigma}$ と Y_i を定義する。

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$$

一方，$\mathrm{E}(Y_i) = 0$, $\mathrm{V}(Y_i) = 1$ を利用して，Y_i の積率母関数は，

$$\begin{aligned}\phi(\theta) &\equiv \mathrm{E}(e^{Y_i\theta}) \\ &= \mathrm{E}\left(1 + Y_i\theta + \frac{1}{2}Y_i^2\theta^2 + \frac{1}{3!}Y_i^3\theta^3 \cdots\right) \\ &= 1 + \frac{1}{2}\theta^2 + O(\theta^3)\end{aligned}$$

となる。($Y_i = 0$ の回りで，$e^{Y_i\theta}$ をテーラー展開する。)
$Z = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{i=1}^{n} Y_i$ の積率母関数 $\Phi(\theta)$ は，

$$\begin{aligned}\Phi(\theta) &= \mathrm{E}(e^{Z\theta}) \\ &= \mathrm{E}\left(e^{\frac{\theta}{\sqrt{n}}\sum_{i=1}^{n} Y_i}\right) \\ &= \prod_{i=1}^{n} \mathrm{E}\left(e^{\frac{\theta}{\sqrt{n}}Y_i}\right) \\ &= \prod_{i=1}^{n} \phi_i\left(\frac{\theta}{\sqrt{n}}\right) \\ &= \left(\phi\left(\frac{\theta}{\sqrt{n}}\right)\right)^n \\ &= \left(1 + \frac{1}{2}\frac{\theta^2}{n} + O(\frac{\theta^3}{n^{\frac{3}{2}}})\right)^n \\ &= \left(1 + \frac{1}{2}\frac{\theta^2}{n} + O(n^{-\frac{3}{2}})\right)^n\end{aligned}$$

注)
$Y_1, Y_2, \cdots, Y_n$ はそれぞれ独立で，同じ分布関数を持つので，積率母関数は同じになる。すなわち，

$$\phi_1(\theta) = \phi_2(\theta) = \ \cdots \ = \phi_n(\theta) = \phi(\theta)$$

となる。
さらに，$x = \dfrac{1}{2}\dfrac{\theta^2}{n} + O(n^{-\frac{3}{2}})$ とおき，$\dfrac{n}{x}$ を両辺に掛けて，$n = \dfrac{1}{x}\Big(\dfrac{1}{2}\theta^2 + O(n^{-\frac{1}{2}})\Big)$ を代入する。

$$\begin{aligned}\Phi(\theta) &= \left(1 + \frac{1}{2}\frac{\theta^2}{n} + O(n^{-\frac{3}{2}})\right)^n \\ &= (1+x)^{\frac{1}{x}(\frac{\theta^2}{2} + O(n^{-\frac{1}{2}}))} \\ &= \left((1+x)^{\frac{1}{x}}\right)^{\frac{\theta^2}{2} + O(n^{-\frac{1}{2}})} \ \longrightarrow \ e^{\frac{\theta^2}{2}}\end{aligned}$$

$n \longrightarrow \infty$ のとき，$x \longrightarrow 0$ となる。
注)
e の定義について，

$$e = \lim_{x \to 0}(1+x)^{\frac{1}{x}} = \lim_{h \to \infty}\left(1 + \frac{1}{h}\right)^h = 2.71828182845905$$

に注意。

$e^{\frac{\theta^2}{2}}$ は標準正規分布 $N(0,1)$ の積率母関数であるので，

$$P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} < x\right) \ \longrightarrow \ \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \mathrm{d}u$$

が成り立つ。

系： $X_1, X_2, \cdots, X_n$ が同じ分布に従わなくても，独立性もなくても，$n \longrightarrow \infty$ のとき，

$$P\left(\frac{\sum_{i=1}^n X_i - \mathrm{E}(\sum_{i=1}^n X_i)}{\sqrt{\mathrm{V}(\sum_{i=1}^n X_i)}} < x\right) \longrightarrow \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \mathrm{d}u$$

が成り立つ。(証明略)

7 大数の強法則 (Almost sure convergence)

大数の強法則 (その 1)： $X_1, X_2, \cdots, X_n$ が互いに独立な確率変数ですべて同じ分布に従うとし，すべての i について $\mathrm{E}(X_i) = \mu$ とする。このとき，

$$P\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1$$

ただし，$\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^n X_i$ とする。

大数の強法則 (その 2)： $X_1, X_2, \cdots, X_n$ が互いに独立な確率変数で，平均と分散が存在するものとする (同一の分布に従わなくてもよい)。$\mathrm{V}(X_i) = \sigma_i^2$, $S_n = \sum_{i=1}^n X_n$, $m_n = \mathrm{E}(S_n)$ とする。$\sum_{i=1}^\infty \dfrac{\sigma_i^2}{i^2} < \infty$ のとき，

$$P\left(\lim_{n\to\infty} \frac{S_n - m_n}{n} = 0\right) = 1$$

が成り立つ。

8 統計的推定

8.1 推定法と標本平均および標本分散の性質

8.1.1 推定法

点推定： 母集団の分布型は既知，その分布のある特性値 θ (母数) は未知とする。
その母集団の分布は $f(x;\theta)$ で与えられている。
このとき，標本の実現値 $(x_1,\ x_2,\ \cdots,\ x_n)$ から適当な値 $\widehat{\theta}_n(x_1, x_2, \cdots, x_n)$ を計算する。
$\widehat{\theta}_n(x_1, x_2, \cdots, x_n)$ を θ の推定値とする。$\Longrightarrow$ 点推定
例：
母平均 μ の点推定値 (＝標本平均 $\overline{x}$)

$$\widehat{\mu}_n(x_1, x_2, \cdots, x_n) \equiv \overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$$

母分散 σ^2 の点推定値 (＝標本不偏分散 s^2)

$$\widehat{\sigma}_n^2(x_1, x_2, \cdots, x_n) \equiv s^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x})^2$$

区間推定： 母集団の分布の未知母数 θ を推定するとき，実現値 $(x_1, x_2, \cdots, x_n)$ より $\widehat{\theta}_L(x_1, x_2, \cdots, x_n)$ と $\widehat{\theta}_U(x_1, x_2, \cdots, x_n)$ を作り，区間 $(\widehat{\theta}_L,\widehat{\theta}_U)$ の中に θ は $1-\alpha$ の確率で入っていることを示す推定法を区間推定法という。
区間 $(\widehat{\theta}_L,\widehat{\theta}_U)$ は信頼係数 $1-\alpha$ の θ の信頼区間である。
$\widehat{\theta}_L \Longrightarrow$ 信頼下限
$\widehat{\theta}_U \Longrightarrow$ 信頼上限
信頼区間の幅はなるべく狭くなるように $\widehat{\theta}_L, \widehat{\theta}_U$ を選ぶ。

8.1.2 標本，統計量，推定量

母集団の分布型は既知，その分布のある特性値 θ (母数) は未知とする。
その母集団の分布は $f(x;\theta)$ で与えられている。
標本： $X_1, X_2, \cdots, X_n \Longrightarrow$ 母集団の部分集合
実現値： $x_1, x_2, \cdots, x_n$
母数 θ の推定量： $\widehat{\theta}(X_1, X_2, \cdots, X_n)$
母数 θ の推定値： $\widehat{\theta}(x_1, x_2, \cdots, x_n)$

例： $\theta = (\mu, \sigma^2)$ とする。
μ の推定量： $\overline{X} = \dfrac{1}{n}\sum_{i=1}^n X_i$
μ の推定値： $\overline{x} = \dfrac{1}{n}\sum_{i=1}^n x_i$
σ^2 の推定量： $S^2 = \dfrac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$
σ^2 の推定値： $s^2 = \dfrac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$
注)
確率変数の関数 $\Longrightarrow$ 統計量
母数の推定のために使われる統計量 $\Longrightarrow$ 推定量

8.1.3　母平均，母分散の推定

$X_1, X_2, \cdots, X_n$ は互いに独立で，すべて同一の分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。

1. 母平均 μ の推定量：
$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
2. 母分散 σ^2 の推定量：
   - 母平均 μ が既知のとき： $S^{*2} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \mu)^2$
   - 母平均 μ が未知のとき： $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$

$\overline{X}$ の性質：

$$\begin{aligned}
\mathrm{E}(\overline{X}) &= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n}\mathrm{E}(\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mu \\
&= \frac{1}{n}n\mu \\
&= \mu
\end{aligned}$$

$$\begin{aligned}
\mathrm{V}(\overline{X}) &= \mathrm{V}(\frac{1}{n}\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n^2}\mathrm{V}(\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}(X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 \\
&= \frac{1}{n^2}n\sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

必要な公式：
$\mathrm{E}(aX) = a\mathrm{E}(X)$
$\mathrm{V}(aX) = a^2\mathrm{V}(X)$
$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$
$\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y) \Longleftarrow X$ と Y は独立のとき

S^{*2}, S^2 の性質：

$$\begin{aligned}
\mathrm{E}(S^{*2}) &= \mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\Big((X_i-\mu)^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{V}(X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sigma^2 \\
&= \frac{1}{n}n\sigma^2 \\
&= \sigma^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}(S^2) &= \mathrm{E}\Big(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\overline{X})^2\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}\big((X_i-\mu)-(\overline{X}-\mu)\big)^2\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}\big((X_i-\mu)^2 \\
&\qquad -2(X_i-\mu)(\overline{X}-\mu) \\
&\qquad +(\overline{X}-\mu)^2\big)\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2 \\
&\qquad -2(\overline{X}-\mu)\sum_{i=1}^{n}(X_i-\mu) \\
&\qquad +n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2 \\
&\qquad -2n(\overline{X}-\mu)^2 + n(\overline{X}-\mu)^2\Big)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2\Big) \\
&\qquad -\frac{1}{n-1}\mathrm{E}\Big(n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{E}\Big((X_i-\mu)^2\Big) \\
&\qquad -\frac{n}{n-1}\mathrm{E}\Big((\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{V}(X_i) - \frac{n}{n-1}\mathrm{V}(\overline{X}) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\sigma^2 - \frac{n}{n-1}\frac{\sigma^2}{n} \\
&= \frac{1}{n-1}n\sigma^2 - \frac{1}{n-1}\sigma^2 \\
&= \sigma^2
\end{aligned}
$$

必要な公式：

$\mathrm{E}(X_i-\mu)^2 = \mathrm{V}(X_i) = \sigma^2$

$\mathrm{E}(\overline{X}-\mu)^2 = \mathrm{V}(\overline{X}) = \dfrac{\sigma^2}{n}$

$\displaystyle\sum_{i=1}^{n}(X_i-\mu) = n(\overline{X}-\mu)$

したがって，$S^{**2} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2$ とすると，

$$
\begin{aligned}
\mathrm{E}(S^{**2}) &= \mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\Big) \\
&= \mathrm{E}\Big(\frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\Big) \\
&= \mathrm{E}\Big(\frac{n-1}{n}S^2\Big) \\
&= \frac{n-1}{n}\mathrm{E}(S^2) \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}
$$

8.2 点推定法： 最適性

母数： θ

推定値： $\widehat{\theta}_n(x_1, x_2, \cdots, x_n)$

推定量： $\widehat{\theta}_n(X_1, X_2, \cdots, X_n)$

$\widehat{\theta}_n(X_1, X_2, \cdots, X_n)$ を $\widehat{\theta}_n$ と書く。

$\widehat{\theta}_n$ の望ましい性質：不偏性，有効性，十分性，一致性
$\Longrightarrow$ 最適性
$\Longrightarrow$ 最適推定量

不偏性：　$\mathrm{E}(\widehat{\theta}_n) = \theta$
$\Longrightarrow$ θ を中心にして $\widehat{\theta}_n$ は分布している。
$\widehat{\theta}_n$ は θ の不偏推定量であるという。
$\mathrm{E}(\widehat{\theta}_n) - \theta$ をバイアス (bias) と呼ぶ。

X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。

1. 母平均 μ の推定量：

   $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$

2. 母分散 σ^2 の推定量：

   $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2$

$\mathrm{E}(\overline{X}) = \mu$, $\mathrm{E}(S^2) = \sigma^2$ なので，$\overline{X}$, S^2 は μ, σ^2 の不偏推定量である。

有効性：　2 つの不偏推定量 $\widehat{\theta}_n$，$\widetilde{\theta}_n$ を考える。
すなわち，$\mathrm{E}(\widehat{\theta}_n) = \theta$, $\mathrm{E}(\widetilde{\theta}_n) = \theta$
$\mathrm{V}(\widehat{\theta}_n) < \mathrm{V}(\widetilde{\theta}_n)$ のとき，$\widehat{\theta}_n$ が $\widetilde{\theta}_n$ より有効であるという。
$\Longrightarrow$ バラツキの小さい推定量の方が望まれる。

クラメール・ラオの不等式 (Cramer-Rao Inequality): 任意の不偏推定量 $\widehat{\theta}_n$ について，

$$\mathrm{V}(\widehat{\theta}_n) \geq \frac{\sigma^2}{n}$$

が成り立つ。ただし，

$$
\begin{aligned}
\sigma^2 &= \sigma^2(\theta) \\
&= \frac{1}{\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right]} \\
&= -\frac{1}{\mathrm{E}\left(\dfrac{\partial^2 \log f(X;\theta)}{\partial\theta^2}\right)}
\end{aligned}
$$

とする。
等号が成り立つ $\widehat{\theta}_n$ が存在するとき，$\widehat{\theta}_n$ は最小分散の不偏推定量である。
$\Longrightarrow$ 有効推定量

クラメール・ラオの不等式の証明：
まず，準備として，尤度関数 $l(\theta;x)=l(\theta;x_1,x_2,\cdots,x_n)$ は $(X_1,X_2,\cdots,X_n)$ の結合密度関数なので，その積分値は 1 となる。
尤度関数については，後述。
すなわち，

$$1=\int l(\theta;x)\mathrm{d}x$$

を得る。
ただし，$l(\theta;x)=\prod_{i=1}^{n}f(x_i;\theta)$ とする。
$\int\cdots\mathrm{d}x$ は n 重積分を意味するものとする。
両辺を θ で微分して整理する。

$$\begin{aligned}0&=\int\frac{\partial l(\theta;x)}{\partial\theta}\mathrm{d}x\\&=\int\frac{1}{l(\theta;x)}\frac{\partial l(\theta;x)}{\partial\theta}l(\theta;x)\mathrm{d}x\\&=\int\frac{\partial\log l(\theta;x)}{\partial\theta}l(\theta;x)\mathrm{d}x\\&=\mathrm{E}\left[\frac{\partial\log l(\theta;X)}{\partial\theta}\right]\end{aligned}$$

これは $\dfrac{\partial\log l(\theta;X)}{\partial\theta}$ の期待値はゼロを意味する。
3 行目では，$\dfrac{\mathrm{d}\log x}{\mathrm{d}x}=\dfrac{1}{x}$ に注意。

今，θ の推定量を $\widehat{\theta}_n$ とおく。

$$\mathrm{E}(\widehat{\theta}_n)=\int\widehat{\theta}_n l(\theta;x)\mathrm{d}x$$

θ について微分

$$\begin{aligned}&\frac{\partial\mathrm{E}(\widehat{\theta}_n)}{\partial\theta}\\&=\int\widehat{\theta}_n\frac{\partial l(\theta;x)}{\partial\theta}\mathrm{d}x\\&=\int\widehat{\theta}_n\frac{\partial\log l(\theta;x)}{\partial\theta}l(\theta;x)\mathrm{d}x\\&=\int\Big(\widehat{\theta}_n-\mathrm{E}(\widehat{\theta}_n)\Big)\Big(\frac{\partial\log l(\theta;x)}{\partial\theta}-\mathrm{E}(\frac{\partial\log l(\theta;x)}{\partial\theta})\Big)\\&\qquad\times l(\theta;x)\mathrm{d}x\\&=\mathrm{Cov}\left(\widehat{\theta}_n,\frac{\partial\log l(\theta;X)}{\partial\theta}\right)\end{aligned}$$

簡単化のために，$\widehat{\theta}_n$，θ をスカラーとする。
このとき，

$$\begin{aligned}&\left(\frac{\partial\mathrm{E}(\widehat{\theta}_n)}{\partial\theta}\right)^2\\&=\left[\mathrm{Cov}\left(\widehat{\theta}_n,\frac{\partial\log l(\theta;X)}{\partial\theta}\right)\right]^2\\&=\rho^2\mathrm{V}\left(\widehat{\theta}_n\right)\mathrm{V}\left(\frac{\partial\log l(\theta;X)}{\partial\theta}\right)\\&\le\mathrm{V}\left(\widehat{\theta}_n\right)\mathrm{V}\left(\frac{\partial\log l(\theta;X)}{\partial\theta}\right)\end{aligned}$$

ただし，ρ は $\widehat{\theta}_n$ と $\dfrac{\partial\log l(\theta;X)}{\partial\theta}$ との相関係数とする。すなわち，$-1\le\rho\le1$ で，その定義は，

$$\rho=\frac{\mathrm{Cov}\left(\widehat{\theta}_n,\dfrac{\partial\log l(\theta;X)}{\partial\theta}\right)}{\sqrt{\mathrm{V}\left(\widehat{\theta}_n\right)}\sqrt{\mathrm{V}\left(\dfrac{\partial\log l(\theta;X)}{\partial\theta}\right)}}$$

となる。よって，

$$\left(\frac{\partial\mathrm{E}(\widehat{\theta}_n)}{\partial\theta}\right)^2\le\mathrm{V}(\widehat{\theta}_n)\ \mathrm{V}\left(\frac{\partial\log l(\theta;X)}{\partial\theta}\right)$$

すなわち，

$$\mathrm{V}(\widehat{\theta}_n)\ge\frac{\left(\dfrac{\partial\mathrm{E}(\widehat{\theta}_n)}{\partial\theta}\right)^2}{\mathrm{V}\left(\dfrac{\partial\log l(\theta;X)}{\partial\theta}\right)}$$

$\mathrm{E}(\widehat{\theta}_n)=\theta$ のとき，

$$\begin{aligned}\mathrm{V}(\widehat{\theta}_n)&\ge\frac{\left(\dfrac{\partial\mathrm{E}(\widehat{\theta}_n)}{\partial\theta}\right)^2}{\mathrm{V}\left(\dfrac{\partial\log l(\theta;X)}{\partial\theta}\right)}\\&=\frac{1}{\mathrm{E}\left[\left(\dfrac{\partial\log l(\theta;X)}{\partial\theta}\right)^2\right]}\end{aligned}$$

さらに，

$$\mathrm{E}\left[\left(\frac{\partial\log l(\theta;X)}{\partial\theta}\right)^2\right]$$

$$
\begin{aligned}
&= \mathrm{E}\left[\left(\sum_{i=1}^{n}\frac{\partial \log f(X_i;\theta)}{\partial\theta}\right)^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n}\left(\frac{\partial \log f(X_i;\theta)}{\partial\theta}\right)^2\right] \\
&= \sum_{i=1}^{n}\mathrm{E}\left[\left(\frac{\partial \log f(X_i;\theta)}{\partial\theta}\right)^2\right] \qquad (\text{同一の分布により}) \\
&= \sum_{i=1}^{n}\mathrm{E}\left[\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right] \\
&= n\mathrm{E}\left[\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right] \\
&= n\int_{-\infty}^{\infty}\left(\frac{\partial \log f(x;\theta)}{\partial\theta}\right)^2 f(x;\theta)\mathrm{d}x
\end{aligned}
$$

X_i, $i=1,2,\cdots,n$, は互いに独立なので，2 つ目の等号が成り立つ。

したがって，

$$
\mathrm{V}(\widehat{\theta}_n) \geq \frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right]} = \frac{\sigma^2(\theta)}{n}
$$

となる。

次に，

$$
\begin{aligned}
&-\mathrm{E}\left(\frac{\partial^2 \log f(X;\theta)}{\partial\theta^2}\right) \\
&= \mathrm{E}\left[\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right] \\
&= \mathrm{V}\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)
\end{aligned}
$$

を証明する。

$$
\int f(x;\theta)\mathrm{d}x = 1
$$

θ について微分

$$
\int \frac{\partial f(x;\theta)}{\partial\theta}\mathrm{d}x = 0
$$

(x の範囲は θ に依存しないもの，微分 $\partial f(x;\theta)/\partial\theta$ が存在するものと仮定される)

上式の変形により

$$
\int \frac{\partial \log f(x;\theta)}{\partial\theta} f(x;\theta)\mathrm{d}x = 0
$$

すなわち，

$$
\mathrm{E}\left(\frac{\partial \log f(x;\theta)}{\partial\theta}\right) = 0
$$

さらに，θ について微分

$$
\begin{aligned}
&\int \frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} f(x;\theta)\mathrm{d}x \\
&+ \int \left(\frac{\partial \log f(x;\theta)}{\partial\theta}\right)^2 \mathrm{d}x \\
&= \int \frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} f(x;\theta)\mathrm{d}x \\
&+ \int \left(\frac{\partial \log f(x;\theta)}{\partial\theta}\right)^2 f(x;\theta)\mathrm{d}x
\end{aligned}
$$

よって，

$$
\begin{aligned}
&-\mathrm{E}\left(\frac{\partial^2 \log f(X;\theta)}{\partial\theta\partial\theta'}\right) \\
&= \mathrm{E}\left[\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right] \\
&= \mathrm{V}\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)
\end{aligned}
$$

を得る。

したがって，

$$
\begin{aligned}
\sigma^2(\theta) &= \frac{1}{\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right]} \\
&= -\frac{1}{\mathrm{E}\left(\dfrac{\partial^2 \log f(X;\theta)}{\partial\theta^2}\right)}
\end{aligned}
$$

となる。

例： X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の正規分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。ただし，σ^2 は既知とする。このとき，μ の推定量 $\overline{X}$ は有効推定量である。

証明：

$\mathrm{V}(\overline{X})$ は分布形にかかわらず，$\sigma^2 < \infty$ のとき，$\dfrac{\sigma^2}{n}$ となる。 .... (A)

一方，

$$
f(x;\mu) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)
$$

なので，クラメール・ラオの不等式は，この場合，

$$\mathrm{V}(\overline{X}) \geq \frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\mu)}{\partial \mu}\right)^2\right]}$$

となる。

$$\log f(X;\mu) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(X-\mu)^2$$

なので，

$$\frac{\partial \log f(X;\mu)}{\partial \mu} = \frac{1}{\sigma^2}(X-\mu)$$

となる。
したがって，

$$\begin{aligned}\mathrm{V}(\overline{X}) &\geq \frac{1}{n\mathrm{E}\left[\left(\dfrac{1}{\sigma^2}(X-\mu)\right)^2\right]} \\ &= \frac{1}{n\dfrac{1}{\sigma^4}\mathrm{E}[(X-\mu)^2]} \\ &= \frac{\sigma^2}{n}\end{aligned}$$

となる。...(B)
(A) と (B) より，$\mathrm{V}(\overline{X}) = \dfrac{\sigma^2}{n}$ (クラメール・ラオの不等式の下限) が成り立つので，$\overline{X}$ は有効推定量であると言える。

十分性 (充分性，充足性)： $X_1, X_2, \cdots, X_n$ の同時密度関数が，

$$\begin{aligned}&f(x_1, x_2, \cdots, x_n;\theta) \\ &= h(x_1, x_2, \cdots, x_n)g\Big(\widehat{\theta}_n(x_1, x_2, \cdots, x_n);\theta\Big)\end{aligned}$$

と分解できるとき，$\widehat{\theta}_n \equiv \widehat{\theta}_n(X_1, X_2, \cdots, X_n)$ は θ の十分推定量であるという。
すなわち，
$h(x_1, x_2, \cdots, x_n;\theta) = h(x_1, x_2, \cdots, x_n)$
が成り立つとき ($\widehat{\theta}_n$ を与えたときの X_1, X_2, $\cdots$, X_n の条件付き分布が母数 θ に依存しないとき)，$\widehat{\theta}_n \equiv \widehat{\theta}_n(X_1, X_2, \cdots, X_n)$ は θ の十分推定量となる。
$\Longrightarrow$ 標本 X_1, X_2, $\cdots$, X_n に含まれている θ に関する情報は，すべて $\widehat{\theta}_n$ に含まれている。

X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の正規分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。ただし，σ^2 は既知とする。このとき，$\overline{X}$ は十分推定量である。
証明：

$$\begin{aligned}&f(x_1, x_2, \cdots, x_n;\mu) \\ &= \prod_{i=1}^{n} f(x_i;\mu) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \\ &\quad\times \exp\left(-\frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n}(x_i-\overline{x})^2 + n(\overline{x}-\mu)^2\Big)\right) \\ &= \frac{1}{\sqrt{n}(\sqrt{2\pi\sigma^2})^{n-1}}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2\right) \\ &\quad\times \frac{1}{\sqrt{2\pi\sigma^2/n}}\exp\left(-\frac{1}{2\sigma^2/n}(\overline{x}-\mu)^2\right) \\ &= h(x_1, x_2, \cdots, x_n)g(\overline{x};\mu)\end{aligned}$$

以上のように，分解可能である。
したがって，$\overline{X}$ は十分推定量である。

十分推定量の補足
Rao-Blackwell の定理：
t を θ の十分統計量，$\widehat{\theta}_n$ を θ の不偏推定量とする。このとき，$\mathrm{E}(\widehat{\theta}_n|t)$ は θ の不偏推定量で，しかも，$\widehat{\theta}_n$ よりも分散は小さい。
$\Longrightarrow$ 任意の不偏推定量 $\widehat{\theta}_n$ について，$\widehat{\theta}_n = v(t)$ でない限り (不偏推定量 $\widehat{\theta}_n$ が t だけの関数でない限り)，$v(t) = \mathrm{E}(\widehat{\theta}_n|t)$ を作ることにより改善される。
どの $\widehat{\theta}_n$ に対しても，同じ $v(t)$ が一意に決まれば，それ以上の改善は出来ないので，$v(t)$ は θ のあらゆる不偏推定量の中で最小な分散を持つことになる。
$\Longrightarrow$ $v(t)$ は有効推定量

一致性： n が大きくなるにつれて，任意の $\epsilon > 0$ について，

$$P(|\widehat{\theta}_n - \theta| > \epsilon) \;\longrightarrow\; 0, \quad n \;\longrightarrow\; \infty$$

となるとき，$\widehat{\theta}_n$ は θ の一致推定量であるという。

X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の正規分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従う

ものとする。ただし，σ^2 は既知とする。このとき，$\overline{X}$ は一致推定量である。

証明：

チェビシェフの不等式：

$$P(|X-\mu|>\epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

ただし，$\mathrm{E}(X)=\mu, \mathrm{V}(X)=\sigma^2$ とする。

ここで，X を $\overline{X}$ にして，$\mathrm{E}(\overline{X}), \mathrm{V}(\overline{X})$ を求める。

$$\mathrm{E}(\overline{X})=\mu, \quad \mathrm{V}(\overline{X})=\frac{\sigma^2}{n}$$

を得る。

したがって，$n \longrightarrow \infty$ のとき，

$$P(|\overline{X}-\mu|>\epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \longrightarrow 0$$

となる。

よって，$\overline{X}$ は μ の一致推定量である。

定理

$$\lim_{n\to\infty} \mathrm{E}(\widehat{\theta}_n)=\theta \qquad \text{(漸近的不偏)}$$

かつ

$$\lim_{n\to\infty} \mathrm{V}(\widehat{\theta}_n)=0$$

のとき，$\widehat{\theta}_n$ は θ の一致推定量となる。

8.3 推定量の求め方： 最尤法, 積率法, 最小二乗法

8.3.1 最尤法

標本 $X_1, X_2, \cdots, X_n$ の密度関数：

$$f(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)$$

θ は未知母数 $\Longrightarrow \widehat{\theta}_n(x_1,x_2,\cdots,x_n)$ によって推定

$$l(\theta)=l(\theta;x)=l(\theta;x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n} f(x_i;\theta)$$

のように，θ の関数として考える。

$l(\theta)$： 尤度関数

尤度関数を最大にする θ を $\widehat{\theta}_n$ とする。

$\widehat{\theta}_n \equiv \widehat{\theta}_n(X_1,X_2,\cdots,X_n) \Longrightarrow$ 最尤推定量

$\widehat{\theta}_n(x_1,x_2,\cdots,x_n) \Longrightarrow$ 最尤推定値

すなわち，

$$\frac{\partial l(\theta)}{\partial \theta}=0$$

を解くことによって，最尤推定量 $\widehat{\theta}_n \equiv \widehat{\theta}_n(X_1,X_2,\cdots,X_n)$ が得られる。

最尤推定量の性質：

小標本について (n が小さいとき)：

- 一般に，最尤推定量は不偏性を持っていないが，適当な変換によって，不偏推定量を作ることが出来る場合が多い。
- 有効推定量が存在すれば (すなわち，クラメール・ラオの不等式の等号を満たすような推定量が存在するならば)，最尤推定量は有効推定量に一致する。
- 十分統計量が存在すれば，最尤推定量は十分統計量の関数となる。

大標本について (n が大きいとき)：

$n \longrightarrow \infty$ のとき，

$$\sqrt{n}(\widehat{\theta}_n-\theta) \longrightarrow N\Big(0,\sigma^2\Big)$$

となる。$\Longrightarrow$ 一致性，漸近的正規性，漸近的有効性

ただし，

$$\sigma^2=\sigma^2(\theta)=\frac{1}{\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\theta)}{\partial \theta}\right)^2\right]}$$

すなわち，$n \longrightarrow \infty$ のとき，

$$\frac{\sqrt{n}(\widehat{\theta}_n-\theta)}{\sigma(\theta)}=\frac{\widehat{\theta}_n-\theta}{\sigma(\theta)/\sqrt{n}} \longrightarrow N(0,1)$$

となる。

したがって，厳密ではないが，n が大きいとき，

$$\widehat{\theta}_n \sim N\left(\theta, \frac{\sigma^2(\theta)}{n}\right)$$

と近似できる。

すなわち，$n \longrightarrow \infty$ のとき，$\widehat{\theta}_n$ の分散はクラメール・ラオの不等式の下限 $\dfrac{\sigma^2(\theta)}{n}$ に近づくことを意味する。

$\Longrightarrow$ 漸近的に有効推定量

さらに，分母の θ を最尤推定量 $\widehat{\theta}_n$ で置き換えて，$n \longrightarrow \infty$ のとき，

$$\frac{\widehat{\theta}_n - \theta}{\sigma(\widehat{\theta}_n)/\sqrt{n}} \longrightarrow N(0,1)$$

となる。
実際には，n が大きいとき，

$$\widehat{\theta}_n \sim N\left(\theta, \frac{\sigma^2(\widehat{\theta}_n)}{n}\right)$$

と近似して用いる。

例：
X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の正規分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。μ, σ^2 の最尤推定量を求める。

$$\begin{aligned}
&f(x_1, x_2, \cdots, x_n; \mu, \sigma^2) \\
&= \prod_{i=1}^{n} f(x_i; \mu, \sigma^2) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right) \\
&= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right) \\
&= l(\mu, \sigma^2)
\end{aligned}$$

対数をとる。(最大化しやすくなる場合が多い)

$$\log l(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

対数尤度関数 $\log l(\mu, \sigma^2)$ を μ と σ^2 について微分して，ゼロと置く。

$$\begin{aligned}
\frac{\partial \log l(\mu, \sigma^2)}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) \\
&= 0 \\
\frac{\partial \log l(\mu, \sigma^2)}{\partial \sigma^2} &= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 \\
&= 0
\end{aligned}$$

この 2 つの連立方程式を解く。

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

μ, σ^2 の最尤推定量は，

$$\overline{X}, \quad S^{**2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

となる。
$\mathrm{E}(\overline{X}) = \mu$ なので，μ の最尤推定量 $\overline{X}$ は不偏推定量である。
$\mathrm{E}(S^{**2}) = \dfrac{n-1}{n}\sigma^2 \neq \sigma^2$ なので，σ^2 の最尤推定量 S^{**2} は不偏推定量でない。

例：
X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一のベルヌイ分布ですべて同一の分布) に従うものとする。すなわち，X の確率関数は $P(X = x) = f(x; p) = p^x(1-p)^{1-p}$, $x = 0, 1$, となる。p の最尤推定量を求める。

$$\begin{aligned}
&f(x_1, x_2, \cdots, x_n; p) \\
&= \prod_{i=1}^{n} f(x_i; p) \\
&= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= p^{\sum_i x_i}(1-p)^{n-\sum_i x_i} \\
&= l(p)
\end{aligned}$$

対数をとる。

$$\log l(p) = (\sum_i x_i)\log(p) + (n - \sum_i x_i)\log(1-p)$$

対数尤度関数 $\log l(p)$ を p について微分して，ゼロと置く。

$$\begin{aligned}
\frac{\partial \log l(p)}{\partial p} &= \frac{\sum_i x_i}{p} - \frac{n - \sum_i x_i}{1-p} \\
&= \frac{\sum_i x_i - np}{p(1-p)} \\
&= 0
\end{aligned}$$

この方程式を解く。

$$p = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

p の最尤推定量は，

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

となる。
$\mathrm{E}(\overline{X}) = p$ なので，p の最尤推定量 $\overline{X}$ は不偏推定量である。
X がベルヌイ分布 $f(x;p)$ のとき，$\mathrm{E}(X) = p$ に注意。

例：

X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一のポアソン分布 (すなわち，平均 λ ですべて同一の分布) に従うものとする。λ の最尤推定量を求める。
ポアソン分布の確率関数は,

$$P(X=x) = f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \cdots$$

なので，尤度関数は,

$$\begin{aligned} l(\lambda) &= \prod_{i=1}^{n} f(x_i;\lambda) \\ &= \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \\ &= \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} \end{aligned}$$

対数尤度関数は,

$$\log l(\lambda) = \log(\lambda)\sum_{i=1}^{n} x_i - n\lambda - \log(\prod_{i=1}^{n} x_i!)$$

となる。

$$\begin{aligned} \frac{\partial \log l(\lambda)}{\partial \lambda} &= \frac{1}{\lambda}\sum_{i=1}^{n} x_i - n \\ &= 0 \end{aligned}$$

これを解いて，λ の最尤推定量 $\widehat{\lambda}$ は,

$$\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

となる。

$\widehat{\lambda}$ は，λ の不偏推定量，有効推定量，十分推定量，一致推定量である。

証明：

X がパラメータ λ のポアソン分布に従うとき,

$$\mathrm{E}(X) = \mathrm{V}(X) = \lambda$$

となる。

不偏性：

$$\begin{aligned} \mathrm{E}(\widehat{\lambda}) &= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} \lambda \\ &= \lambda \end{aligned}$$

有効性：

$$\begin{aligned} \mathrm{V}(\widehat{\lambda}) &= \mathrm{V}(\frac{1}{n}\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{V}(X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \lambda \\ &= \frac{\lambda}{n} \end{aligned}$$

$$\begin{aligned} &\frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\lambda)}{\partial \lambda}\right)^2\right]} \\ &= \frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial (X\log\lambda - \lambda - \log X!)}{\partial \lambda}\right)^2\right]} \\ &= \frac{1}{n\mathrm{E}\left[\left(\dfrac{X}{\lambda} - 1\right)^2\right]} \\ &= \frac{\lambda^2}{n\mathrm{E}[(X-\lambda)^2]} \\ &= \frac{\lambda^2}{n\mathrm{V}(X)} \\ &= \frac{\lambda^2}{n\lambda} \\ &= \frac{\lambda}{n} \end{aligned}$$

したがって,

$$\mathrm{V}(\widehat{\lambda}) = \frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\lambda)}{\partial \lambda}\right)^2\right]}$$

となり，$\mathrm{V}(\widehat{\lambda})$ は，クラメール・ラオの下限に一致する。よって，$\widehat{\lambda}$ は有効推定量である。

十分性：

$$\begin{aligned}\prod_{i=1}^{n} f(x_i;\lambda) &= \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} \\ &= \frac{\lambda^{n\overline{x}} e^{-n\lambda}}{(n\overline{x})!} \frac{(n\overline{x})!}{\prod_{i=1}^{n} x_i!} \\ &= g(\overline{x};\lambda)\ h(x_1, x_2, \cdots, x_n)\end{aligned}$$

と分解できる。
一致性：

$$\mathrm{E}(\overline{X}) = \lambda, \qquad \mathrm{V}(\overline{X}) = \frac{\lambda}{n}$$

なので，チェビシェフの不等式に当てはめる。

$$P(|\overline{X} - \lambda| > \epsilon) < \frac{\lambda}{n\epsilon^2} \ \longrightarrow\ \infty$$

したがって，一致性も成り立つ。

8.3.2 積率法 (モーメント法)

ν 次の原点の回りの積率： $\mathrm{E}(X^\nu) \equiv \mu'_\nu = \mu'_\nu(\theta)$
$\mathrm{E}(X^\nu)$ の推定量を $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^\nu$ とする。

$$\frac{1}{n}\sum_{i=1}^{n} X_i^\nu = \mu'_\nu(\widehat{\theta}_1, \widehat{\theta}_2, \cdots, \widehat{\theta}_k)$$

$\nu = 1, 2, \cdots, k$ として，k 個の連立方程式を $\widehat{\theta}_1, \widehat{\theta}_2, \cdots, \widehat{\theta}_k$ について解く。
その解：
$\widehat{\theta}_1(X_1, X_2, \cdots, X_n)$,
$\widehat{\theta}_2(X_1, X_2, \cdots, X_n)$,
$\vdots$
$\widehat{\theta}_k(X_1, X_2, \cdots, X_n)$
$\Longrightarrow$ 積率法 (モーメント法) による推定量

例：
X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。μ, σ^2 の積率法による推定値を求める。

$$\mathrm{E}(X) = \mu \ \Longrightarrow\ \frac{1}{n}\sum_{i=1}^{n} X_i = \widehat{\mu}$$

$$\mathrm{E}(X^2) = \sigma^2 + \mu^2 \ \Longrightarrow\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \widehat{\sigma}^2 + \widehat{\mu}^2$$

連立方程式を解いて，

$$\widehat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

を得る。この場合，最尤推定量に一致する。

8.3.3 最小二乗法

X_i と μ の差 $X_i - \mu$ を考える。
$X_i - \mu$ の二乗和を最小にする μ を求める。
すなわち，

$$\min_{\mu}\sum_{i=1}^{n}(X_i - \mu)^2$$

を最小にする μ を求める。$\Longrightarrow$

$$\frac{\mathrm{d}S(\mu)}{\mathrm{d}\mu} = 0$$

の解を求める。ただし，

$$S(\mu) = \sum_{i=1}^{n}(X_i - \mu)^2$$

とする。

$$\frac{\mathrm{d}S(\mu)}{\mathrm{d}\mu} = -2\sum_{i=1}^{n}(X_i - \mu) = 0$$

を満たす μ を $\widehat{\mu}$ とする。$\Longrightarrow$

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

を得る。$\Longrightarrow$ 最小二乗推定量

9 標本分布

9.1 正規母集団の場合 (標本平均，標本不偏分散の標本分布)

大きさ n の無作為標本 $\{X_1,\, X_2,\, \cdots,\, X_n\}$。
すべての $i = 1, 2, \cdots, n$ について，$X_i \sim N(\mu, \sigma^2)$ とする。

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

9.1.1 正規分布：標本平均 $\overline{X}$ の標本分布

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

証明：
積率母関数を利用する。$X \sim N(\mu, \sigma^2)$ のとき，X の積率母関数 $\phi(\theta)$ は，

$$\begin{aligned}
\phi(\theta) &\equiv \mathrm{E}(e^{\theta X}) \\
&= \int_{-\infty}^{\infty} e^{\theta x} f(x)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2+\theta x}\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x^2-2(\mu+\sigma^2\theta)x+\mu^2\right)}\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x-(\mu+\sigma^2\theta)\right)^2+(\mu\theta+\frac{1}{2}\sigma^2\theta^2)}\mathrm{d}x \\
&= e^{\mu\theta+\frac{1}{2}\sigma^2\theta^2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x-(\mu+\sigma^2\theta)\right)^2}\mathrm{d}x \\
&= \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right)
\end{aligned}$$

と計算される。積分のところは，$N(\mu+\sigma^2\theta, \sigma^2)$ の確率密度関数に注意
よって，$X_i \sim N(\mu, \sigma^2)$ のとき，X_i の積率母関数 $\phi_i(\theta)$ は，

$$\phi_i(\theta) = \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right)$$

となる。
今，$\overline{X}$ の積率母関数 $\phi_{\overline{x}}(\theta)$ を考える。

$$\begin{aligned}
\phi_{\overline{x}}(\theta) &\equiv \mathrm{E}(e^{\theta\overline{X}}) \\
&= \mathrm{E}(e^{\theta\frac{1}{n}\sum_{i=1}^{n}X_i}) \\
&= \mathrm{E}(\prod_{i=1}^{n} e^{\frac{\theta}{n}X_i}) \\
&= \prod_{i=1}^{n}\mathrm{E}(e^{\frac{\theta}{n}X_i}) \\
&= \prod_{i=1}^{n}\phi_i(\frac{\theta}{n}) \\
&= \prod_{i=1}^{n}\exp\left(\mu\frac{\theta}{n} + \frac{1}{2}\sigma^2\left(\frac{\theta}{n}\right)^2\right) \\
&= \exp\left(\mu\theta + \frac{1}{2}\sigma^2\frac{\theta^2}{n}\right) \\
&= \exp\left(\mu\theta + \frac{1}{2}\frac{\sigma^2}{n}\theta^2\right)
\end{aligned}$$

となり，これは，平均 μ，分散 σ^2/n の正規分布の積率母関数に一致する。

さらに，標準化によって，

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

を得る。これも，同様に，積率母関数で証明できる。

9.1.2 χ^2（カイ自乗）分布：標本不偏分散 S^2 の標本分布

$$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$$

証明：

$$\frac{X_i-\mu}{\sigma} \sim N(0,1)$$

なので，

$$\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(1)$$

$\dfrac{X_i-\mu}{\sigma}$, $i = 1, 2, \cdots, n$ は互いに独立なので，

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$$

となる。(積率母関数によって証明可)

μ を $\overline{X}$ で置き換えることによって，

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

を得る。(証明略)

9.1.3 t 分布：標本平均 $\overline{X}$ の標本分布

$Z \sim N(0,1)$, $U \sim \chi^2(m)$, Z と U は独立のとき，

$$\frac{Z}{\sqrt{U/m}} \sim t(m)$$

となる。

これを利用して，$\overline{X}$ の標本分布を求める。

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ と $\dfrac{(n-1)S^2}{\sigma^2}$ は独立なので (後述),

$$\frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/(n-1)}} = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

を得る。

(*) $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ と $\dfrac{(n-1)S^2}{\sigma^2}$ の独立性

証明：

X_1, X_2, $\cdots$, X_n は互いに独立に正規分布に従うので，$\mathrm{E}\big((\overline{X}-\mu)S^2\big)=0$ を証明すればよい。

$$\overline{X}-\mu = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)$$

$$\begin{aligned} S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\Big((X_i-\mu)-(\overline{X}-\mu)\Big)^2 \\ &= \frac{1}{n-1}\Big(\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2\Big) \end{aligned}$$

$$\begin{aligned} &\mathrm{E}\big((\overline{X}-\mu)S^2\big) \\ &= \frac{1}{n-1}\mathrm{E}\Big((\overline{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^3\Big) \\ &= \frac{n}{n-1}\mathrm{E}\big(\sum_{j=1}^{n}(X_j-\mu)\sum_{i=1}^{n}(X_i-\mu)^2\big) \\ &\quad - \frac{n}{n-1}\mathrm{E}\big((\overline{X}-\mu)^3\big) \\ &= \frac{n}{n-1}\mathrm{E}\big(\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i-\mu)^2(X_j-\mu)\big) \\ &\quad - \frac{n}{n-1}\mathrm{E}\big((\overline{X}-\mu)^3\big) \\ &= \frac{n}{n-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}\big((X_i-\mu)^2(X_j-\mu)\big) \\ &\quad - \frac{n}{n-1}\mathrm{E}\big((\overline{X}-\mu)^3\big) \\ &= \frac{n}{n-1}\sum_{i=1}^{n}\mathrm{E}\big((X_i-\mu)^3\big) \\ &\quad - \frac{n}{n-1}\mathrm{E}\big((\overline{X}-\mu)^3\big) \\ &= 0 \end{aligned}$$

$Z \sim N(\mu_z, \sigma_z^2)$ のとき，$\mathrm{E}\big((Z-\mu_z)^{2k-1}\big)=0,\, k=1,2,\cdots$ となる。(積率母関数より証明可)

したがって，$\mathrm{E}\big((X_i-\mu)^3\big)=\mathrm{E}\big((\overline{X}-\mu)^3\big)=0$ となる。

よって，

$$\mathrm{E}\big((\overline{X}-\mu)S^2\big)=0$$

を得る。

9.1.4 F 分布

$U\sim\chi^2(n)$, $V\sim\chi^2(m)$, U と V は独立のとき，

$$\frac{U/n}{V/m} \sim F(n,m)$$

となる。

公式：

$t(m)^2 = F(1,m)$

$F(n,m) = 1/F(m,n)$

9.2 その他の母集団の場合：標本平均 $\overline{X}$ の標本分布

中心極限定理を利用する。

中心極限定理： 大きさ n の無作為標本 X_1, X_2, $\cdots$, X_n すべての i について，$\mathrm{E}(X_i)=\mu$, $\mathrm{V}(X_i)=\sigma^2$ とする。

標本平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1^n X_i}$ の分布を考える。

$n \longrightarrow \infty$ のとき，

$$\begin{aligned} \frac{\overline{X}-\mathrm{E}(\overline{X})}{\sqrt{\mathrm{V}(\overline{X})}} &= \frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}} \\ &= \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \longrightarrow N(0,1) \end{aligned}$$

が成り立つ。
さらに，分母の σ をその標本不偏分散 S で置き換えても，

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \longrightarrow N(0,1)$$

が成り立つ。(証明略)

例題： 母比率の推定量の標本分布： 離散型確率変数 X の取りうる値は 0, 1 のどちらかで，その確率分布は,

$$P(X=x)=f(x)=p^x(1-p)^{1-x} \qquad x=0,1$$

となる。すなわち，

$$\begin{aligned} P(X=1)&=p \\ P(X=0)&=1-p \end{aligned}$$

となる。
何等かの実験を行い，成功すれば $X=1$，失敗すれば $X=0$ として，確率変数 X に数値を割り当てる。
または，アンケート調査によって，Yes と答えれば $X=1$，No と答えれば $X=0$ として，確率変数 X に数値を割り当てる。
このとき，$X=1$ となる確率 $P(X=1)=p$ を推定することを考える。
$\mathrm{E}(X)$, $\mathrm{V}(X)$ を求めておく。

$$\begin{aligned} \mathrm{E}(X)&=\sum_x xf(x) \\ &=\sum_{x=0}^{1} xp^x(1-p)^{1-x} \\ &=p \end{aligned}$$

$\mu=\mathrm{E}(X)$ のとき，$\mathrm{V}(X)=\mathrm{E}(X^2)-\mu^2$ により，まず，$\mathrm{E}(X^2)$ を求める。

$$\begin{aligned} \mathrm{E}(X^2)&=\sum_x x^2f(x) \\ &=\sum_{x=0}^{1} x^2p^x(1-p)^{1-x} \\ &=p^2 \end{aligned}$$

よって，

$$\begin{aligned} \mathrm{V}(X)&=\mathrm{E}(X^2)-\mu^2 \\ &=p-p^2 \\ &=p(1-p) \end{aligned}$$

n 個の無作為標本 $X_1, X_2, \cdots, X_n$ を考える。
$X_1, X_2, \cdots, X_n$ はそれぞれ独立に，しかも上記のように，同一のベルヌイ分布に従うものとする。
このときの母比率 p の推定量を $\widehat{P}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$ とする。
すなわち，$\widehat{P}$ は n 人の中で成功した回数 (または，Yes と答えた人数) を表し，標本平均である。
中心極限定理を当てはめる。$n \longrightarrow \infty$ のとき，

$$\frac{\widehat{P}-\mathrm{E}(\widehat{P})}{\sqrt{\mathrm{V}(\widehat{P})}}=\frac{\widehat{P}-p}{\sqrt{p(1-p)/n}} \longrightarrow N(0,1)$$

が成り立つ。$\mathrm{E}(\widehat{P})$, $\mathrm{V}(\widehat{P})$ は

$$\begin{aligned} \mathrm{E}(\widehat{P})&=\mathrm{E}(\frac{1}{n}\sum_{i=1}^{n}X_i) \\ &=\frac{1}{n}\mathrm{E}(\sum_{i=1}^{n}X_i) \\ &=\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) \\ &=\frac{1}{n}\sum_{i=1}^{n}p \\ &=p \\ \mathrm{V}(\widehat{P})&=\mathrm{V}(\frac{1}{n}\sum_{i=1}^{n}X_i) \\ &=\frac{1}{n^2}\mathrm{V}(\sum_{i=1}^{n}X_i) \\ &=\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}(X_i) \longleftarrow \text{互いに独立} \\ &=\frac{1}{n^2}\sum_{i=1}^{n}p(1-p) \\ &=\frac{p(1-p)}{n} \end{aligned}$$

となることに注意。$\mathrm{E}(X_i)=p$, $\mathrm{V}(X_i)=p(1-p)$ を途中で利用する。
さらに，分母の p をその推定量 $\widehat{P}$ で置き換えても，

$$\frac{\widehat{P}-p}{\sqrt{\widehat{P}(1-\widehat{P})/n}} \longrightarrow N(0,1)$$

が成り立つ。(証明略)

9.3 その他の母集団の場合： 母数の推定量 $\widehat{\theta}_n$ の標本分布 (一般化)

母集団の分布： $f(x;\theta)$
より一般的に，母数 θ の推定量の標本分布を求める。
最尤推定量の性質を利用する。
すなわち，母数 θ の最尤推定量 $\widehat{\theta}_n$ の標本分布について：
$n \longrightarrow \infty$ のとき，

$$\sqrt{n}(\widehat{\theta}_n - \theta) \longrightarrow N\left(0, \sigma^2\right)$$

となる。ただし，

$$\sigma^2 = \sigma^2(\theta) = \left(\mathrm{E}\left[\left(\frac{\partial \log f(X;\theta)}{\partial \theta}\right)^2\right]\right)^{-1}$$

となる。
標準化 (基準化) によって，$n \longrightarrow \infty$ のとき，

$$\frac{\sqrt{n}(\widehat{\theta}_n - \theta)}{\sigma(\theta)} = \frac{\widehat{\theta}_n - \theta}{\sigma(\theta)/\sqrt{n}} \longrightarrow N(0,1)$$

となる。
さらに，分母の θ を最尤推定量 $\widehat{\theta}_n$ で置き換えても，$n \longrightarrow \infty$ のとき，

$$\frac{\widehat{\theta}_n - \theta}{\sigma(\widehat{\theta}_n)/\sqrt{n}} \longrightarrow N(0,1)$$

が成り立つ。(証明略)

10 区間推定法

母集団の分布が $f(x;\theta)$ で表されているとき，この母集団からの標本 $X_1, X_2, \cdots, X_n$ を使って，θ に依存しない 2 つの統計量 $\widehat{\theta}_U(X_1, X_2, \cdots, X_n)$, $\widehat{\theta}_L(X_1, X_2, \cdots, X_n)$ を作り，

$$P\Big(\widehat{\theta}_L(X_1, X_2, \cdots, X_n) < \theta < \widehat{\theta}_U(X_1, X_2, \cdots, X_n)\Big) = 1-\alpha$$

と表すことが出来たとき，
区間 $\Big(\widehat{\theta}_L(x_1, x_2, \cdots, x_n), \widehat{\theta}_U(x_1, x_2, \cdots, x_n)\Big)$
を信頼係数 $1-\alpha$ (または，有意水準 α) の信頼区間と呼ぶ。
$\widehat{\theta}_L(x_1, x_2, \cdots, x_n), \widehat{\theta}_U(x_1, x_2, \cdots, x_n)$ をそれぞれ信頼下限，信頼上限という。
一般に，
信頼区間の幅 $\widehat{\theta}_U(X_1, X_2, \cdots, X_n) - \widehat{\theta}_L(X_1, X_2, \cdots, X_n)$ が最小になることが望ましい。

10.1 母平均 μ の区間推定

10.1.1 正規母集団の場合 (小標本，大標本共に)

大きさ n の無作為標本 $\{X_1, X_2, \cdots, X_n\}$。すべての $i = 1, 2, \cdots, n$ について，$X_i \sim N(\mu, \sigma^2)$ とする。

母分散 σ^2 が既知のとき：

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

$$P\Big(\Big|\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\Big| < z_{\alpha/2}\Big) = 1-\alpha$$

$z_{\alpha/2}$ は $100 \times \dfrac{\alpha}{2}$ % 点で，確率 α が与えられると，正規分布表から得られる。
したがって，$P\Big(\overline{X} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\Big) = 1-\alpha$
$\overline{X}$ を $\overline{x}$ で置き換えて，信頼係数 $1-\alpha$ の μ の信頼区間：
$\Longrightarrow \boxed{\Big(\overline{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \overline{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\Big)}$

母分散 σ^2 が未知のとき：

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

$$P\Big(\Big|\frac{\overline{X} - \mu}{S/\sqrt{n}}\Big| < t_{\alpha/2}(n-1)\Big) = 1-\alpha$$

$t_{\alpha/2}(n-1)$ は $100 \times \dfrac{\alpha}{2}$ % 点で，確率 α と自由度 $n-1$ が与えられると，t 分布表から得られる。
したがって，$P\Big(\overline{X} - t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}\Big) = 1-\alpha$
$\overline{X}$, S^2 を $\overline{x}$, s^2 で置き換えて，信頼係数 $1-\alpha$ の μ の信頼区間：$\Longrightarrow \boxed{\Big(\overline{x} - t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}, \overline{x} + t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}\Big)}$

10.1.2 その他の母集団の場合：大標本 (n が大きいとき)

大きさ n の無作為標本 $\{X_1, X_2, \cdots, X_n\}$
すべての $i = 1, 2, \cdots, n$ について，$X_i \sim (\mu, \sigma^2)$ とする (正規分布を仮定する必要ない)。

母分散 σ^2 が既知のとき： n が大きいとき，中心極限定理により，以下が成り立つ。

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

$$P\Big(\Big|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\Big| < z_{\alpha/2}\Big) \approx 1-\alpha$$

したがって，$P\Big(\overline{X}-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} < \mu < \overline{X}+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\Big) \approx 1-\alpha$
$\overline{X}$ を $\overline{x}$ で置き換えて，信頼係数 $1-\alpha$ の μ の信頼区間：
$\Longrightarrow \boxed{\Big(\overline{x}-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \overline{x}+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\Big)}$

母分散 σ^2 が未知のとき： さらに，分母の σ^2 を標本不偏分散 S^2 で置き換えて，近似する。

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1),$$

$$P\Big(\Big|\frac{\overline{X}-\mu}{S/\sqrt{n}}\Big| < z_{\alpha/2}\Big) \approx 1-\alpha$$

したがって，$P\Big(\overline{X}-z_{\alpha/2}\dfrac{S}{\sqrt{n}} < \mu < \overline{X}+z_{\alpha/2}\dfrac{S}{\sqrt{n}}\Big) \approx 1-\alpha$
$\overline{X}$, S^2 を $\overline{x}$, s^2 で置き換えて，信頼係数 $1-\alpha$ の μ の信頼区間：$\Longrightarrow \boxed{\Big(\overline{x}-z_{\alpha/2}\dfrac{s}{\sqrt{n}}, \overline{x}+z_{\alpha/2}\dfrac{s}{\sqrt{n}}\Big)}$

10.2 母分散 σ^2 の区間推定 (正規母集団)

大きさ n の無作為標本 $\{X_1, X_2, \cdots, X_n\}$
すべての $i=1,2,\cdots,n$ について，$X_i \sim N(\mu,\sigma^2)$ とする

$$\frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$$

μ をその推定量 $\overline{X}$ で置き換えると，

$$U = \sum_{i=1}^n \Big(\frac{X_i-\overline{X}}{\sigma}\Big)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

ただし，$S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (X_i-\overline{X})^2$ (標本不偏分散) とする。

1. 自由度 $n-1$ のカイ 2 乗分布の下側確率，上側確率が $\alpha/2$ となる点を，それぞれ $\chi^2_{1-\alpha/2}(n-1)$，$\chi^2_{\alpha/2}(n-1)$ とする。
2. $\chi^2_{1-\alpha/2}(n-1)$
   $\Longrightarrow$ 下側確率が $\alpha/2$ となる点
   $\Longrightarrow$ 上側確率が $1-\alpha/2$ となる点
3. $P\big(\chi^2_{1-\alpha/2}(n-1) < U < \chi^2_{\alpha/2}(n-1)\big) = 1-\alpha$
   $P\big(\chi^2_{1-\alpha/2}(n-1) < \dfrac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}(n-1)\big) = 1-\alpha$
   $P\big(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\big) = 1-\alpha$
4. 推定量 S^2 をその推定値 s^2 で置き換えて，
   信頼係数 $1-\alpha$ の σ^2 の信頼区間は，
   $\boxed{\Big(\dfrac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\Big)}$
   となる。
   ただし，$s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (x_i-\overline{x})^2$ とする。

10.3 母比率 p の区間推定 (ベルヌイ試行)

p は次の確率を表す。

$$P(X=1)=p$$
$$P(X=0)=1-p$$

すなわち，何等かの実験を行い，成功すれば $X=1$，失敗すれば $X=0$ として，確率変数 X に数値を割り当てる。
または，アンケート調査によって，Yes と答えれば $X=1$，No と答えれば $X=0$ として，確率変数 X に数値を割り当てる。
このとき，$X=1$ となる確率 $P(X=1)=p$ を推定することを考える。
n 個の無作為標本 $X_1, X_2, \cdots, X_n$ を考える。
$X_1, X_2, \cdots, X_n$ はそれぞれ独立に，しかも上記のように，同一のベルヌイ分布に従うものとする。
母比率 p の推定量は $\widehat{P} = \dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i$ とする。
すなわち，$\widehat{P}$ は n 人の中で成功した回数 (または，Yes と答えた人数) を表し，標本平均である。
$\mathrm{E}(\widehat{P})=p$, $\mathrm{V}(\widehat{P})=p(1-p)/n$ を用いて，中心極限定理により，n が大きいとき，近似的に，

$$\frac{\widehat{P}-p}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

を得る。
さらに，分母の p をその推定量 $\widehat{P}$ で置き換えて，近似する。

$$\frac{\widehat{P}-p}{\sqrt{\widehat{P}(1-\widehat{P})/n}} \sim N(0,1),$$

$$P\Bigl(\left|\frac{\widehat{P}-p}{\sqrt{\widehat{P}(1-\widehat{P})/n}}\right| < z_{\alpha/2}\Bigr) \approx 1-\alpha$$

したがって，

$P\Bigl(\widehat{P}-z_{\alpha/2}\sqrt{\dfrac{\widehat{P}(1-\widehat{P})}{n}} < p < \widehat{P}+z_{\alpha/2}\sqrt{\dfrac{\widehat{P}(1-\widehat{P})}{n}}\Bigr) \approx 1-\alpha$

$\widehat{P}$ を $\widehat{p}$ で置き換えて，信頼係数 $1-\alpha$ の p の信頼区間：

$\Longrightarrow \boxed{\Bigl(\widehat{p}-z_{\alpha/2}\sqrt{\dfrac{\widehat{p}(1-\widehat{p})}{n}}, \widehat{p}+z_{\alpha/2}\sqrt{\dfrac{\widehat{p}(1-\widehat{p})}{n}}\Bigr)}$

11　統計的検定 I

11.1　仮説検定の考え方

仮説検定：母集団の分布 $f(x;\theta)$ が与えられているときに，母数 θ についての仮説 $\theta=\theta_0$ が正しいかどうかを，標本 (X_1, X_2, $\cdots$, X_n) の実現値 (x_1, x_2, $\cdots$, x_n) から判断する。
帰無仮説 H_0：$\theta=\theta_0$

2 種類の誤り： 検定しようとする仮説 $\Longrightarrow$ 帰無仮説 H_0
帰無仮説が正しくないときに成り立つ仮説 $\Longrightarrow$ 対立仮説 H_1

	H_0 は正しい	H_0 は正しくない
H_0 採択	正しい判定	第 2 種の誤り (確率 β)
H_0 棄却	第 1 種の誤り (確率 $\alpha=$ 有意水準)	正しい判定 ($1-\beta=$ 検出力)

検定の手続き

1. 母数について帰無仮説 H_0 を立てる。
2. ある適当な統計量を考えて，H_0 が正しいときにこの統計量が従う分布を導く。

   ある適当な統計量 $\Longrightarrow$ 検定統計量
3. 実際の標本 (データ) からこの統計量の値 (統計値) を計算する。
4. 統計量の分布と統計値とを比較する。

   この統計値が分布の端にあれば，H_0 は起こりにくいと判定され，H_0 を棄却する。

起こりにくいとして H_0 を棄却する領域 $\Longrightarrow$ 棄却域 R (reject の意味)
起こり得るとして H_0 を採択する領域 $\Longrightarrow$ 採択域 A (accept の意味)

検定統計量 $T=f(X_1,X_2,\cdots,X_n)$
第 1 種の誤りの確率 $=$ 有意水準 α (H_0 が正しいにもかかわらず，H_0 を棄却する確率)

$$P\bigl(f(X_1,X_2,\cdots,X_n)\in R|H_0 \text{ が正しい}\bigr)=\alpha$$

第 2 種の誤りの確率 $=\beta$ (H_0 が正しくないにもかかわらず，H_0 を採択する確率)

$$P\bigl(f(X_1,X_2,\cdots,X_n)\in A|H_0 \text{ が正しくない}\bigr)=\beta$$

検出力 $=1-\beta$ (H_0 が正しくないとき，H_0 を棄却する確率)

$$P\bigl(f(X_1,X_2,\cdots,X_n)\in R|H_0 \text{ が正しくない}\bigr)=1-\beta$$

有意水準：$\alpha=0.05, 0.01$ を選ぶ。

検出力の導出方法：正規母集団で，分散が既知の場合： X_1, X_2, $\cdots$, X_n の n 個の確率変数は，それぞれ独立に，平均 μ，分散 σ^2 の正規分布をするものとする。ただし，分散 σ^2 は既知とする。
帰無仮説 H_0：$\mu=\mu_0$，対立仮説 H_1：$\mu=\mu_1$ の検定を考える。$\mu_1>\mu_0$ とする。
検出力を求める。
検出力とは，対立仮説のもとで，帰無仮説を棄却する確率である。
標本平均 $\overline{X}\sim N(\mu,\dfrac{\sigma^2}{n})$ から，

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \ \sim\ N(0,1)$$

で，帰無仮説 H_0：$\mu=\mu_0$ のもとで，

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \ \sim\ N(0,1)$$

となる。
よって，帰無仮説 H_0： $\mu=\mu_0$ のもとで，帰無仮説を棄却する棄却域は，$P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}>z_\alpha\Big)=\alpha$ なので，
$\overline{X}>\mu_0+z_\alpha\sigma/\sqrt{n}$ となる。
対立仮説 H_1： $\mu=\mu_1$ のもとで，帰無仮説を棄却する確率 (検出力) を求める。
すなわち，対立仮説 H_1： $\mu=\mu_1$ のもとで，
$P\Big(\overline{X}>\mu_0+z_\alpha\sigma/\sqrt{n}\Big)$ の確率を求めればよい。
対立仮説 H_1： $\mu=\mu_1$ のもとで，
$\dfrac{\overline{X}-\mu_1}{\sigma/\sqrt{n}}\sim N(0,1)$ となるので，
$P\Big(\dfrac{\overline{X}-\mu_1}{\sigma/\sqrt{n}}>\dfrac{\mu_0-\mu_1}{\sigma/\sqrt{n}}+z_\alpha\Big)$
となる確率が検出力となる。($\dfrac{\mu_0-\mu_1}{\sigma/\sqrt{n}}+z_\alpha$ に対応する確率を標準正規分布表で調べる。)

検定の種類:

1. 母平均 μ の検定
   (a) 正規母集団の場合 (小標本，大標本共に)
      i. 母分散 σ^2 は既知 $\Longrightarrow N(0,1)$
      ii. 母分散 σ^2 は未知 $\Longrightarrow t(n-1)$
   (b) その他の母集団の場合： 大標本 (n が大きいとき) $\Longrightarrow N(0,1)$
2. 2 つの標本の母平均の差 (H_0： $\mu_1=\mu_2$) の検定
   (a) 小標本 (n_1, n_2 が小さいとき)
      i. σ_1^2 と σ_2^2 は既知で，正規母集団の場合 $\Longrightarrow N(0,1)$
      ii. σ_1^2 と σ_2^2 は未知 (一般的に)$\Longrightarrow$ 検定不可能 (検定統計量の分布を導出できないため)
   (b) その他の母集団： 大標本 (n_1, n_2 が大きいとき) $\Longrightarrow N(0,1)$
3. 母分散 σ^2 の検定で，正規母集団の場合 $\Longrightarrow \chi^2(n-1)$
4. 2 つの標本の母分散が等しいかどうか (H_0： $\sigma_1^2=\sigma_2^2$) の検定で，正規母集団の場合 $\Longrightarrow F(n_1-1,n_2-1)$
5. 母比率 p の検定 (ベルヌイ試行) $\Longrightarrow N(0,1)$

11.2　母平均 μ の検定

11.2.1　正規母集団の場合 (小標本，大標本共に)

大きさ n の無作為標本 $\{X_1, X_2, \cdots, X_n\}$。$i=1,2,\cdots,n$ について，$X_i\sim N(\mu,\sigma^2)$ とする。

母分散 σ^2 は既知のとき： $\overline{X}$ の分布は，$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ なので，帰無仮説 H_0： $\mu=\mu_0$ が正しいもとで，
$\boxed{\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)}$ となる (μ を μ_0 で置き換える)。このとき，検定統計量 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$。検定統計量の値 $\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu<\mu_0}$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<-z_\alpha\Big)=\alpha$ なので，$\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}<-z_\alpha$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。
2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu>\mu_0}$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}>z_\alpha\Big)=\alpha$ なので，$\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}>z_\alpha$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。
3. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu\neq\mu_0}$** (両側検定)
   $P\Big(\Big|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\Big|>z_{\alpha/2}\Big)=\alpha$ なので，$\Big|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\Big|>z_{\alpha/2}$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

母分散 σ^2 は未知のとき： $\overline{X}$ の分布は，$\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$ なので，帰無仮説 H_0 ： $\mu=\mu_0$ が正しいもとで，
$\boxed{\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)}$ となる (μ を μ_0 で置き換える)。このとき，検定統計量 $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$。検定統計量の値 $\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu<\mu_0}$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}<-t_\alpha(n-1)\Big)=\alpha$ なので，$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}<-t_\alpha(n-1)$ のとき，有意水準 α で H_0： $\mu=\mu_0$ を棄却する。
2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu>\mu_0}$** (片側検定)

$P\Big(\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} > t_\alpha(n-1)\Big) = \alpha$ なので，$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} > t_\alpha(n-1)$ のとき，有意水準 α で H_0： $\mu=\mu_0$ を棄却する。

3. **対立仮説 H_1： $\mu \neq \mu_0$** (両側検定)
   $P\Big(\Big|\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\Big| > t_{\alpha/2}(n-1)\Big) = \alpha$ なので，$\Big|\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}\Big| > t_{\alpha/2}(n-1)$ のとき，有意水準 α で H_0： $\mu=\mu_0$ を棄却する。

11.2.2 その他の母集団の場合：大標本 (n が大きいとき)

大きさ n の無作為標本 $\{X_1, X_2, \cdots, X_n\}$。$i=1,2,\cdots,n$ について，$X_i \sim (\mu, \sigma^2)$ とする (正規分布の仮定は不必要)。

母分散 σ^2 は既知のとき： 近似的に $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ が成り立つので，帰無仮説 H_0： $\mu=\mu_0$ が正しいもとで，$\boxed{\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)}$ となる (μ を μ_0 で置き換える)。このとき，検定統計量 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$。検定統計量の値 $\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$。

1. **対立仮説 H_1： $\mu < \mu_0$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} < -z_\alpha\Big) \approx \alpha$ なので，$\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} < -z_\alpha$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

2. **対立仮説 H_1： $\mu > \mu_0$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha\Big) \approx \alpha$ なので，$\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

3. **対立仮説 H_1： $\mu \neq \mu_0$** (両側検定)
   $P\Big(\Big|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\Big| > z_{\alpha/2}\Big) \approx \alpha$ なので，$\Big|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\Big| > z_{\alpha/2}$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

母分散 σ^2 は未知のとき： 近似的に，$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ が成り立つので，帰無仮説 H_0： $\mu=\mu_0$ が正しいもとで，$\boxed{\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim N(0,1)}$ となる (μ を μ_0 で置き換える)。このとき，検定統計量 $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$。検定統計量の値 $\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}$。

1. **対立仮説 H_1： $\mu < \mu_0$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} < -z_\alpha\Big) \approx \alpha$ なので，$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} < -z_\alpha$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

2. **対立仮説 H_1： $\mu > \mu_0$** (片側検定)
   $P\Big(\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} > z_\alpha\Big) \approx \alpha$ なので，$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} > z_\alpha$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

3. **対立仮説 H_1： $\mu \neq \mu_0$** (両側検定)
   $P\Big(\Big|\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\Big| > z_{\alpha/2}\Big) \approx \alpha$ なので，$\Big|\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}\Big| > z_{\alpha/2}$ のとき，有意水準 α で帰無仮説 H_0： $\mu=\mu_0$ を棄却する。

11.3 2 つの標本の母平均の差の検定

H_0： $\mu_1 = \mu_2$ の検定

11.3.1 小標本 (n_1, n_2 が小さいとき)

・第 1 グループ：大きさ n_1 の無作為標本。$i=1,2,\cdots,n$ について，$X_{1i} \sim N(\mu_1, \sigma_1^2)$ とする。
・第 2 グループ：大きさ n_2 の無作為標本。$i=1,2,\cdots,n$ について，$X_{2i} \sim N(\mu_2, \sigma_2^2)$ とする。

σ_1^2 と σ_2^2 は既知で，正規母集団の場合： 母平均の差を検定したいので，統計量 $\overline{X}_1 - \overline{X}_2$ の分布を考える。
$\mathrm{E}(\overline{X}_1 - \overline{X}_2) = \mu_1 - \mu_2$,
また，$\mathrm{V}(\overline{X}_1 - \overline{X}_2) = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$ を得る。
したがって，$\overline{X}_1 - \overline{X}_2 \sim N(\mu_1-\mu_2, \dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2})$ が成り立つ (証明略)。さらに，標準化によって，$\dfrac{(\overline{X}_1 - \overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1)$ を得るので，帰無仮説 H_0： $\mu_1=\mu_2$ が正しいもとで，$\boxed{\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1)}$ となる ($\mu_1-\mu_2=0$ を代入する)。このとき，検定統計量 $\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$。検定統計量の値 $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 < \mu_2}$** (片側検定)

   $P\Big(\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} < -z_\alpha\Big) \approx \alpha$ なので，

   $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} < -z_\alpha$ のとき，有意水準 α で H_0 を棄却する。

2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 > \mu_2}$** (片側検定)

   $P\Big(\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} > z_\alpha\Big) \approx \alpha$ なので，

   $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} > z_\alpha$ のとき，有意水準 α で H_0 を棄却する。

3. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 \neq \mu_2}$** (両側検定)

   $P\Big(\Big|\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\Big| > z_{\alpha/2}\Big) \approx \alpha$ なので，

   $\Big|\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\Big| > z_{\alpha/2}$ のとき，有意水準 α で H_0 を棄却する。

$\boldsymbol{\sigma_1^2}$ と $\boldsymbol{\sigma_2^2}$ は未知のとき： 検定統計量の分布を導出できないため，検定不可能 (正規分布や t 分布にはならない)

11.3.2 その他の母集団： 大標本 ($\boldsymbol{n_1}$, $\boldsymbol{n_2}$ が共に大きいとき)

•第 1 グループ：大きさ n_1 の無作為標本。$i=1,2,\cdots,n$ について，$X_{1i}\sim(\mu_1,\sigma_1^2)$ とする (正規分布の仮定は不必要)。
•第 2 グループ：大きさ n_2 の無作為標本。$i=1,2,\cdots,n$ について，$X_{2i}\sim(\mu_2,\sigma_2^2)$ とする (正規分布の仮定は不必要)。

$\boldsymbol{\sigma_1^2}$ と $\boldsymbol{\sigma_2^2}$ は既知のとき： 近似的に，
$\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1)$ を得るので，
帰無仮説 H_0： $\mu_1=\mu_2$ が正しいもとで，
$\boxed{\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1)}$ となる ($\mu_1-\mu_2=0$ を代入する)。

このとき，検定統計量 $\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$。検定統計量の値 $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 < \mu_2}$** (片側検定)

   $P\Big(\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} < -z_\alpha\Big) \approx \alpha$ なので，

   $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} < -z_\alpha$ のとき，有意水準 α で H_0 を棄却する。

2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 > \mu_2}$** (片側検定)

   $P\Big(\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} > z_\alpha\Big) \approx \alpha$ なので，

   $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} > z_\alpha$ のとき，有意水準 α で H_0 を棄却する。

3. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 \neq \mu_2}$** (両側検定)

   $P\Big(\Big|\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\Big| > z_{\alpha/2}\Big) \approx \alpha$ なので，

   $\Big|\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\Big| > z_{\alpha/2}$ のとき，有意水準 α で H_0 を棄却する。

$\boldsymbol{\sigma_1^2}$ と $\boldsymbol{\sigma_2^2}$ は未知のとき： 近似的に，
$\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_1^2/n_1+S_2^2/n_2}} \sim N(0,1)$ を得るので，
帰無仮説 H_0： $\mu_1=\mu_2$ が正しいもとで，
$\boxed{\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}} \sim N(0,1)}$ となる ($\mu_1-\mu_2=0$ を代入する)。
このとき，検定統計量 $\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}}$。検定統計量の値 $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{s_1^2/n_1+s_2^2/n_2}}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1 < \mu_2}$** (片側検定)

$P\Big(\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}}<-z_\alpha\Big)\approx\alpha$ なので，

$\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{s_1^2/n_1+s_2^2/n_2}}<-z_\alpha$ のとき，有意水準 α で H_0 を棄却する。

2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1>\mu_2}$** (片側検定)

   $P\Big(\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}}>z_\alpha\Big)\approx\alpha$ なので，

   $\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{s_1^2/n_1+s_2^2/n_2}}>z_\alpha$ のとき，有意水準 α で H_0 を棄却する。

3. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\mu_1\neq\mu_2}$** (両側検定)

   $P\Big(\Big|\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\Big|>z_{\alpha/2}\Big)\approx\alpha$ なので，

   $\Big|\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{s_1^2/n_1+s_2^2/n_2}}\Big|>z_{\alpha/2}$ のとき，有意水準 α で H_0 を棄却する。

11.4　母分散 σ^2 の検定で，正規母集団の場合

$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ を得るので，帰無仮説 H_0： $\sigma^2=\sigma_0^2$ が正しいもとで， $\boxed{\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)}$ となる (σ^2 を σ_0^2 で置き換える)。このとき，検定統計量 $\dfrac{(n-1)S^2}{\sigma_0^2}$。検定統計量の値 $\dfrac{(n-1)s^2}{\sigma_0^2}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\sigma^2<\sigma_0^2}$** (片側検定)

   $P\Big(\dfrac{(n-1)S^2}{\sigma_0^2}<\chi^2_{1-\alpha}(n-1)\Big)=\alpha$ なので，

   $\dfrac{(n-1)s^2}{\sigma_0^2}<\chi^2_{1-\alpha}(n-1)$ のとき，有意水準 α で H_0 : $\sigma^2=\sigma_0^2$ を棄却する。

2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\sigma^2>\sigma_0^2}$** (片側検定)

   $P\Big(\dfrac{(n-1)S^2}{\sigma_0^2}>\chi^2_{\alpha}(n-1)\Big)=\alpha$ なので，

   $\dfrac{(n-1)s^2}{\sigma_0^2}>\chi^2_{\alpha}(n-1)$ のとき，有意水準 α で H_0 : $\sigma^2=\sigma_0^2$ を棄却する。

3. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\sigma^2\neq\sigma_0^2}$** (両側検定)

   $P\Big(\chi^2_{1-\alpha/2}(n-1)<\dfrac{(n-1)S^2}{\sigma_0^2}<\chi^2_{\alpha/2}(n-1)\Big)=\alpha$ なので，

   $\chi^2_{1-\alpha/2}(n-1)<\dfrac{(n-1)s^2}{\sigma_0^2}<\chi^2_{\alpha/2}(n-1)$ のとき，有意水準 α で H_0 を棄却する。

11.5　2 つの標本の母分散が等しいかどうかの検定で，正規母集団の場合

$\boldsymbol{H_0}$： $\boldsymbol{\sigma_1^2=\sigma_2^2}$ の検定

・第 1 グループ：大きさ n_1 の無作為標本。$i=1,2,\cdots,n$ について，$X_{1i}\sim N(\mu_1,\sigma_1^2)$ とする。
・第 2 グループ：大きさ n_2 の無作為標本。$i=1,2,\cdots,n$ について，$X_{2i}\sim N(\mu_2,\sigma_2^2)$ とする。

$j=1,2$ について，

$$S_j=\frac{1}{n_j-1}\sum_{i=1}^{n_j}(X_{ji}-\overline{X}_j)^2,\quad \overline{X}_j=\frac{1}{n_j}\sum_{i=1}^{n_j}X_{ji}$$

を定義する。

$\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$ を得るので，帰無仮説 H_0 : $\sigma_1^2=\sigma_2^2$ が正しいもとで， $\boxed{\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)}$ となる。このとき，検定統計量 $\dfrac{S_1^2}{S_2^2}$。検定統計量の値 $\dfrac{s_1^2}{s_2^2}$。

1. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\sigma_1^2<\sigma_2^2}$** (片側検定)

   $P\Big(\dfrac{S_1^2}{S_2^2}<F_{1-\alpha}(n_1-1,n_2-1)\Big)=\alpha$ なので，

   $\dfrac{s_1^2}{s_2^2}<F_{1-\alpha}(n_1-1,n_2-1)$ のとき，有意水準 α で H_0： $\sigma_1^2=\sigma_2^2$ を棄却する。

2. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\sigma_1^2>\sigma_2^2}$** (片側検定)

   $P\Big(\dfrac{S_1^2}{S_2^2}>F_{\alpha}(n_1-1,n_2-1)\Big)=\alpha$ なので，

   $\dfrac{s_1^2}{s_2^2}>F_{\alpha}(n_1-1,n_2-1)$ のとき，有意水準 α で H_0 : $\sigma_1^2=\sigma_2^2$ を棄却する。

3. **対立仮説 $\boldsymbol{H_1}$： $\boldsymbol{\sigma^2\neq\sigma_0^2}$** (両側検定)

$P\Big(F_{1-\alpha/2}(n_1-1,n_2-1) < \dfrac{S_1^2}{S_2^2} < F_{\alpha/2}(n_1-1,n_2-1)\Big) = \alpha$ なので,

$F_{1-\alpha/2}(n_1-1,n_2-1) < \dfrac{s_1^2}{s_2^2} < F_{\alpha/2}(n_1-1,n_2-1)$ のとき，有意水準 α で H_0 を棄却する。

$F_{1-\alpha}(n,m) = \dfrac{1}{F_\alpha(m,n)}$ に注意

11.6 母比率 p の検定 (ベルヌイ試行)

中心極限定理により，近似的に，$\dfrac{\widehat{P}-p}{\sqrt{p(1-p)/n}} \sim N(0,1)$ を得るので，帰無仮説 H_0 : $p = p_0$ が正しいもとで，$\boxed{\dfrac{\widehat{P}-p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)}$ となる (p を p_0 で置き換える)。このとき，検定統計量 $\dfrac{\widehat{P}-p_0}{\sqrt{p_0(1-p_0)/n}}$。検定統計量の値 $\dfrac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$。

1. **対立仮説 H_1： $\mu < \mu_0$** (片側検定)

   $P\Big(\dfrac{\widehat{P}-p_0}{\sqrt{p_0(1-p_0)/n}} < -z_\alpha\Big) \approx \alpha$ なので,

   $\dfrac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} < -z_\alpha$ のとき，有意水準 α で $H_0 : \mu = \mu_0$ を棄却する。

2. **対立仮説 H_1： $\mu > \mu_0$** (片側検定)

   $P\Big(\dfrac{\widehat{P}-p_0}{\sqrt{p_0(1-p_0)/n}} > z_\alpha\Big) \approx \alpha$ なので,

   $\dfrac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} > z_\alpha$ のとき，有意水準 α で $H_0 : \mu = \mu_0$ を棄却する。

3. **対立仮説 H_1： $\mu \neq \mu_0$** (両側検定)

   $P\Big(\Big|\dfrac{\widehat{P}-p_0}{\sqrt{p_0(1-p_0)/n}}\Big| > z_{\alpha/2}\Big) \approx \alpha$ なので,

   $\Big|\dfrac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}\Big| > z_{\alpha/2}$ のとき，有意水準 α で H_0 を棄却する。

12 統計的検定 II： 大標本検定

12.1 ワルド (Wald) 検定

最尤推定量の性質を利用
母集団の分布： $f(x;\theta)$
母数 θ の最尤推定量 $\widehat{\theta}_n$ の標本分布について：
$n \longrightarrow \infty$ のとき,

$$\sqrt{n}(\widehat{\theta}_n - \theta) \longrightarrow N(0,\sigma^2)$$

ただし,

$$\begin{aligned}\sigma^2 &= \sigma^2(\theta)\\ &= \Big(\mathrm{E}\Big[\Big(\frac{\partial \log f(X;\theta)}{\partial\theta}\Big)^2\Big]\Big)^{-1}\\ &= -\Big(\mathrm{E}\Big(\frac{\partial^2 \log f(X;\theta)}{\partial\theta^2}\Big)\Big)^{-1}\end{aligned}$$

となる。
$\dfrac{\sigma^2}{n}$ はクラメール・ラオの下限に一致することに注意。

> (復習) クラメール・ラオの不等式：
> 母数 θ の不偏推定量 $\widehat{\theta}_n$ について,
>
> $$\mathrm{V}(\widehat{\theta}_n) \geq \frac{\sigma^2}{n}$$
>
> が成り立つ。σ^2 は上で定義されたものとなる。

言い換えると,

$$\frac{\widehat{\theta}_n - \theta}{\sigma/\sqrt{n}} \longrightarrow N(0,1)$$

を得る。
さらに，分母の θ を最尤推定量 $\widehat{\theta}_n$ で置き換えても，$n \longrightarrow \infty$ のとき，近似的に,

$$\frac{\widehat{\theta}_n - \theta}{\sigma(\widehat{\theta}_n)/\sqrt{n}} \sim N(0,1)$$

が成り立つ。
したがって，H_0： $\theta = \theta_0$ が正しいもとで，$n \longrightarrow \infty$ のとき，近似的に,

$$\frac{\widehat{\theta}_n - \theta_0}{\sigma(\widehat{\theta}_n)/\sqrt{n}} \sim N(0,1)$$

が成り立つ。
よって，$H_0:\ \theta=\theta_0$ と $H_1:\ \theta\neq\theta_0$ について，
$\widehat{\theta}_n$ を実現値で置き換えて，

$$\left|\frac{\widehat{\theta}_n-\theta_0}{\sigma(\widehat{\theta}_n)/\sqrt{n}}\right|>z_{\alpha/2}\quad \text{のとき，} H_0 \text{ を棄却する。}$$

(*) 検定を行う段階では，$\widehat{\theta}_n$ を最尤推定値とみなしていることに注意。すなわち，$\widehat{\theta}_n$ はもともと最尤推定量で $X_1, X_2, \cdots, X_n$ の関数である。しかし，検定を行う段階では，$X_1, X_2, \cdots, X_n$ を $x_1, x_2, \cdots, x_n$ で置き換えて，$\widehat{\theta}_n$ を最尤推定値とみなしている。
また，$H_0:\ \theta=\theta_0$ と $H_1:\ \theta>\theta_0$ について，
$\widehat{\theta}_n$ を実現値で置き換えて，

$$\frac{\widehat{\theta}_n-\theta_0}{\sigma(\widehat{\theta}_n)/\sqrt{n}}>z_{\alpha}\quad \text{のとき，} H_0 \text{ を棄却する。}$$

最後に，$H_0:\ \theta=\theta_0$ と $H_1:\ \theta<\theta_0$ について，
$\widehat{\theta}_n$ を実現値で置き換えて，

$$\frac{\widehat{\theta}_n-\theta_0}{\sigma(\widehat{\theta}_n)/\sqrt{n}}<-z_{\alpha}\quad \text{のとき，} H_0 \text{ を棄却する。}$$

例題： 指数分布から生成された n 個の互いに独立な確率変数 $X_1, X_2, \cdots, X_n$ を考える。指数分布は，

$$f(x;\gamma)=\gamma e^{-\gamma x}\qquad\qquad x>0$$

である。
帰無仮説 $H_0:\ \gamma=\gamma_0$，対立仮説 $H_1:\ \gamma\neq\gamma_0$ を，ワルド検定によって，検定する。

一般的に，母数 γ の最尤推定量 $\widehat{\gamma}_n$ の標本分布について，
$n\longrightarrow\infty$ のとき，近似的に，

$$\frac{\widehat{\gamma}_n-\gamma}{\sigma(\widehat{\gamma}_n)/\sqrt{n}}\ \sim\ N(0,1)$$

が成り立つ。ただし，

$$\begin{aligned}\sigma^2&=\sigma^2(\gamma)\\&=\left(\mathrm{E}\left[\left(\frac{\mathrm{d}\log f(X;\gamma)}{\mathrm{d}\gamma}\right)^2\right]\right)^{-1}\\&=-\left(\mathrm{E}\left(\frac{\mathrm{d}^2\log f(X;\gamma)}{\mathrm{d}\gamma^2}\right)\right)^{-1}\end{aligned}$$

となる。

したがって，$H_0:\ \gamma=\gamma_0$ が正しいもとで，$n\longrightarrow\infty$ のとき，近似的に，

$$\frac{\widehat{\gamma}_n-\gamma_0}{\sigma(\widehat{\gamma}_n)/\sqrt{n}}\ \sim\ N(0,1)$$

が成り立つ。
よって，$H_0:\ \gamma=\gamma_0$ と $H_1:\ \gamma\neq\gamma_0$ について，
$\widehat{\gamma}_n$ を実現値で置き換えて，

$$\left|\frac{\widehat{\gamma}_n-\gamma_0}{\sigma(\widehat{\gamma}_n)/\sqrt{n}}\right|>z_{\alpha/2}\quad \text{のとき，} H_0 \text{ を棄却する。}$$

まず，$\sigma^2(\gamma)$ を求める。

$$\begin{aligned}\sigma^2&=\sigma^2(\gamma)\\&=-\left(\mathrm{E}\left(\frac{\mathrm{d}^2\log f(X;\gamma)}{\mathrm{d}\gamma^2}\right)\right)^{-1}\\&=\gamma^2\end{aligned}$$

$\log f(X;\gamma)$ の 2 回微分は，

$$\begin{aligned}\frac{\mathrm{d}\log f(X;\gamma)}{\mathrm{d}\gamma}&=\frac{1}{\gamma}-X\\\frac{\mathrm{d}^2\log f(X;\gamma)}{\mathrm{d}\gamma^2}&=-\frac{1}{\gamma^2}\end{aligned}$$

に注意。

次に，γ の最尤推定量 $\widehat{\gamma}_n$ を求める。
$X_1,\cdots,X_n$ は互いに独立で、指数分布 $f(x;\gamma)$ に従うので，尤度関数 $l(\gamma)$ は

$$\begin{aligned}l(\gamma)&=\prod_{i=1}^{n}f(x_i;\gamma)\\&=\prod_{i=1}^{n}\gamma e^{-\gamma x_i}\\&=\gamma^n e^{-\gamma\sum x_i}\end{aligned}$$

対数尤度関数は

$$\log l(\gamma)=n\log(\gamma)-\gamma\sum_{i=1}^{n}x_i$$

$\log l(\gamma)$ を最大にするような γ を求める。

$$\frac{\mathrm{d}\log l(\gamma)}{\mathrm{d}\gamma}=\frac{n}{\gamma}-\sum_{i=1}^{n}x_i=0$$

を解いて，x_i を X_i で置き換えて，γ の最尤推定量 $\widehat{\gamma}_n$ は，

$$\widehat{\gamma}_n=\frac{n}{\sum_{i=1}^{n}X_i}$$

となる。

したがって，$n \longrightarrow \infty$ のとき，近似的に，

$$\frac{\widehat{\gamma}_n - \gamma}{\sigma(\widehat{\gamma}_n)/\sqrt{n}} = \frac{\widehat{\gamma}_n - \gamma}{\widehat{\gamma}_n/\sqrt{n}}$$

が成り立つ。
よって，$H_0:\ \gamma = \gamma_0$ と $H_1:\ \gamma \neq \gamma_0$ について，
$\widehat{\gamma}_n$ を実現値で置き換えて，

$$\left|\frac{\widehat{\gamma}_n - \gamma_0}{\widehat{\gamma}_n/\sqrt{n}}\right| > z_{\alpha/2} \quad \text{のとき，} H_0 \text{ を棄却する。}$$

ただし，γ の最尤推定値 $\widehat{\gamma}_n$ は，

$$\widehat{\gamma}_n = \frac{n}{\sum_{i=1}^n x_i}$$

となる。

12.2 尤度比検定

仮説検定： 母集団の分布 $f(x;\theta)$ が与えられているときに，母数 $\theta = (\theta_1, \theta_2)$ についての仮説 $\theta_1 = \theta_1^*$ が正しいかどうかを，標本 $(X_1, X_2, \cdots, X_n)$ の実現値 $(x_1, x_2, \cdots, x_n)$ から判断する。
θ_1: $1 \times k_1$ ベクトル
θ_2: $1 \times k_2$ ベクトル
$\theta = (\theta_1, \theta_2)$: $1 \times (k_1 + k_2)$ ベクトル
帰無仮説 $H_0:\ \theta_1 = \theta_1^* \Longrightarrow$ 制約の数は k_1 個
尤度関数

$$l(\theta_1, \theta_2) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2)$$

$(\widetilde{\theta}_1, \widetilde{\theta}_2)$ を (θ_1, θ_2) の最尤推定量とする。
すなわち，$(\widetilde{\theta}_1, \widetilde{\theta}_2)$ は，

$$\frac{\partial l(\theta_1, \theta_2)}{\partial \theta_1} = 0, \quad \frac{\partial l(\theta_1, \theta_2)}{\partial \theta_2} = 0$$

の連立方程式を，(θ_1, θ_2) について解いた解である。
$\Longrightarrow$ 制約なし最尤推定量

$\widehat{\theta}_2$ を，帰無仮説 $H_0:\ \theta_1 = \theta_1^*$ が正しいという条件のもとで，θ_2 の最尤推定量とする。
すなわち，$\widehat{\theta}_2)$ は，

$$\frac{\partial l(\theta_1^*, \theta_2)}{\partial \theta_2} = 0$$

を，θ_2 について解いた解である。$\Longrightarrow$ 制約付き最尤推定量 ($\theta_1 = \theta_1^*$ という制約)

尤度比： $\lambda = \dfrac{l(\theta_1^*, \widehat{\theta}_2)}{l(\widetilde{\theta}_1, \widetilde{\theta}_2)}$

n が大きいとき，近似的に，

$$-2\log(\lambda) \sim \chi^2(k_1)$$

となる。「$k_1 =$ 制約の数」に注意。
$-2\log(\lambda) > \chi^2_\alpha(k_1)$ のとき，有意水準 α で，帰無仮説 $H_0:\ \theta_1 = \theta_1^*$ を棄却する。
$-2\log(\lambda)$ がゼロに近ければ，帰無仮説を採択する。
$\Longrightarrow (\theta_1^*, \widehat{\theta}_2)$ が $(\widetilde{\theta}_1, \widetilde{\theta}_2)$ に近い値であれば，$-2\log(\lambda)$ はゼロに近くなる。

母数の推定量 $(\widetilde{\theta}_1, \widetilde{\theta}_2)$ の分布が求めることが出来ない場合に，尤度比検定は有効である。

例題： 指数分布から生成された n 個の互いに独立な確率変数 $X_1, X_2, \cdots, X_n$ を考える。指数分布は，

$$f(x) = \gamma e^{-\gamma x} \qquad x > 0$$

である。
帰無仮説 $H_0:\ \gamma = \gamma_0$ を，尤度比検定によって，検定する。
尤度比

$$\lambda = \frac{l(\gamma_0)}{\max_{\gamma} l(\gamma)} = \frac{l(\gamma_0)}{l(\widehat{\gamma})}$$

について，制約の数は 1 なので，

$$-2\log\lambda \longrightarrow \chi^2(1)$$

となる。
$X_1, \cdots, X_n$ は互いに独立で、指数分布 $f(x)$ に従うので，尤度関数 $l(\gamma)$ は

$$\begin{aligned} l(\gamma) &= \prod_{i=1}^n f(x_i) \\ &= \prod_{i=1}^n \gamma e^{-\gamma x_i} \\ &= \gamma^n e^{-\gamma \sum x_i} \end{aligned}$$

分子について：

$$l(\gamma_0) = \gamma_0^n e^{-\gamma_0 \sum X_i}$$

分母について：
対数尤度関数は

$$\log l(\gamma) = n\log(\gamma) - \gamma\sum_{i=1}^{n} x_i$$

$\log l(\gamma)$ を最大にするような γ を求める。

$$\frac{\mathrm{d}\log l(\gamma)}{\mathrm{d}\gamma} = \frac{n}{\gamma} - \sum_{i=1}^{n} x_i = 0$$

を解いて，x_i を X_i で置き換えて，γ の最尤推定量 $\widehat{\gamma}$ は

$$\widehat{\gamma} = \frac{n}{\sum_{i=1}^{n} X_i}$$

となる。

$$l(\widehat{\gamma}) = \widehat{\gamma}^n e^{-n}$$

したがって，尤度比

$$\begin{aligned}\lambda &= \frac{l(\gamma_0)}{l(\widehat{\gamma})}\\ &= \frac{\gamma_0^n e^{-\gamma_0 \sum X_i}}{\widehat{\gamma}^n e^{-n}}\end{aligned}$$

を得る。
よって，近似的に，

$$-2\log\lambda \sim \chi^2(1)$$

となる。
$-2\log\lambda > \chi^2_\alpha(1)$ のとき，有意水準 α で，帰無仮説 H_0 : $\mu = \mu_0$ を棄却する。$\chi^2_\alpha(1)$ は自由度 1 のカイ二乗分布の $100\times\alpha$ % 点とする。

例題： $X_1, X_2, \cdots, X_n$ の n 個の確率変数は，それぞれ独立に，平均 μ，分散 σ^2 の分布をするものとする。
平均 μ，分散 σ^2 の正規分布の密度関数は

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

で与えられる。
帰無仮説 H_0 : $\mu = \mu_0$ を，尤度比検定によって，検定する。

尤度比

$$\lambda = \frac{\max_{\sigma^2} l(\mu_0,\sigma^2)}{\max_{\mu,\sigma^2} l(\mu,\sigma^2)} = \frac{l(\mu_0,\widetilde{\sigma}^2)}{l(\widehat{\mu},\widehat{\sigma}^2)}$$

について，制約の数は 1 なので，

$$-2\log\lambda \longrightarrow \chi^2(1)$$

となる。
ここで，$l(\mu,\sigma^2)$ は，

$$\begin{aligned}&l(\mu,\sigma^2)\\ &= f(x_1,x_2,\cdots,x_n;\mu,\sigma^2)\\ &= \prod_{i=1}^{n} f(x_i;\mu,\sigma^2)\\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\\ &= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\end{aligned}$$

であり，$\log l(\mu,\sigma^2)$ は，

$$\begin{aligned}\log l(\mu,\sigma^2) = &-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2)\\ &- \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\end{aligned}$$

である。

分子について：
$\mu = \mu_0$ のもとで，$\log l(\mu_0,\sigma^2)$ を σ^2 について最大化する。

$$\frac{\partial \log l(\mu_0,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu_0)^2 = 0$$

この解が $\widetilde{\sigma}^2$ である。したがって，

$$\widetilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2$$

を得る。よって，$l(\mu_0,\widetilde{\sigma}^2)$ は，

$$\begin{aligned}&l(\mu_0,\widetilde{\sigma}^2)\\ &= (2\pi\widetilde{\sigma}^2)^{-n/2} \exp\left(-\frac{1}{2\widetilde{\sigma}^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right)\\ &= (2\pi\widetilde{\sigma}^2)^{-n/2} \exp\left(-\frac{n}{2}\right)\end{aligned}$$

となる。

分母について：

$$\widehat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i$$
$$\widehat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\widehat{\mu})^2$$

を得る。よって，$l(\widehat{\mu},\widehat{\sigma}^2)$ は,

$$\begin{aligned}&l(\widehat{\mu},\widehat{\sigma}^2)\\&=(2\pi\widehat{\sigma}^2)^{-n/2}\exp\left(-\frac{1}{2\widehat{\sigma}^2}\sum_{i=1}^{n}(x_i-\widehat{\mu})^2\right)\\&=(2\pi\widehat{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)\end{aligned}$$

となる。

したがって,

$$\begin{aligned}\lambda&=\frac{\max_{\sigma^2}l(\mu_0,\sigma^2)}{\max_{\mu,\sigma^2}l(\mu,\sigma^2)}\\&=\frac{l(\mu_0,\widetilde{\sigma}^2)}{l(\widehat{\mu},\widehat{\sigma}^2)}\\&=\frac{(2\pi\widetilde{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)}{(2\pi\widehat{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)}\\&=\left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\right)^{-n/2}\end{aligned}$$

よって，近似的に,

$$-2\log\lambda=n(\log\widetilde{\sigma}^2-\log\widehat{\sigma}^2)\sim\chi^2(1)$$

となる。
$-2\log\lambda>\chi^2_\alpha(1)$ のとき，有意水準 α で，帰無仮説 H_0 : $\mu=\mu_0$ を棄却する。$\chi^2_\alpha(1)$ は自由度 1 のカイ二乗分布の $100\times\alpha$ % 点とする。

練習問題と解答 (6 章〜 12 章)

$\boxed{1}$ 平均 μ，分散 σ^2 の正規分布の密度関数は

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

で与えられる。X_1, X_2, $\cdots$, X_n は互いに独立で，しかも，それぞれは平均 μ，分散 σ^2 の正規分布に従うものとする。このとき，次の問に答えよ。

(1) 平均 μ，分散 σ^2 の最尤推定量を求めよ。
(2) σ^2 の最尤推定量は不偏推定量であるかどうかを調べよ。もし，σ^2 の最尤推定量が不偏推定量でないときは，不偏推定量を求めよ (最尤推定量をもとにして考えればよい)。
(3) 帰無仮説 H_0： $\mu=\mu_0$ を，尤度比検定によって，検定したい。どのようにすればよいかを説明せよ。

[解答]

(1) 平均 μ，分散 σ^2 の最尤推定量を求める。

$$\begin{aligned}&f(x_1,x_2,\cdots,x_n;\mu,\sigma^2)\\&=\prod_{i=1}^{n}f(x_i;\mu,\sigma^2)\\&=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\\&=(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\\&=l(\mu,\sigma^2)\end{aligned}$$

対数をとる。(最大化しやすくなる場合が多い)

$$\begin{aligned}\log l(\mu,\sigma^2)=&-\frac{n}{2}\log(2\pi)-\frac{n}{2}\log(\sigma^2)\\&-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\end{aligned}$$

対数尤度関数 $\log l(\mu,\sigma^2)$ を μ と σ^2 について微分して，ゼロと置く。

$$\begin{aligned}\frac{\partial\log l(\mu,\sigma^2)}{\partial\mu}&=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)\\&=0\\\frac{\partial\log l(\mu,\sigma^2)}{\partial\sigma^2}&=-\frac{n}{2}\frac{1}{\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2\\&=0\end{aligned}$$

この 2 つの連立方程式を解く。

$$\mu=\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{x}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$$

μ, σ^2 の最尤推定量は,

$$\overline{X}, \quad S^{**2} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

となる。

(2) σ^2 の最尤推定量 S^{**2} は不偏推定量であるかどうかを調べる。

$$\begin{aligned}
\mathrm{E}(S^{**2}) &= \mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\overline{X})^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}\big((X_i-\mu)-(\overline{X}-\mu)\big)^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}\big((X_i-\mu)^2 \\
&\qquad -2(X_i-\mu)(\overline{X}-\mu) \\
&\qquad +(\overline{X}-\mu)^2\big)\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2 \\
&\qquad -2(\overline{X}-\mu)\sum_{i=1}^{n}(X_i-\mu) \\
&\qquad +n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2 \\
&\qquad -2n(\overline{X}-\mu)^2+n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2\Big) \\
&\qquad -\frac{1}{n}\mathrm{E}\Big(n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\Big((X_i-\mu)^2\Big) \\
&\qquad -\mathrm{E}\Big((\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{V}(X_i)-\mathrm{V}(\overline{X}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sigma^2-\frac{\sigma^2}{n} \\
&= \sigma^2-\frac{1}{n}\sigma^2 \\
&= \frac{n-1}{n}\sigma^2 \\
&\neq \sigma^2
\end{aligned}$$

なので，S^{**2} は σ^2 の不偏推定量ではない。

S^{**2} をもとにして，σ^2 の不偏推定量を求める。

$$\mathrm{E}(S^{**2}) = \frac{n-1}{n}\sigma^2$$

なので，両辺に $\dfrac{n}{n-1}$ をかけて，

$$\frac{n}{n-1}\mathrm{E}(S^{**2}) = \sigma^2$$

を得る。

よって，

$$\frac{n}{n-1}S^{**2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 = S^2$$

が σ^2 の不偏推定量となる。

(3) 帰無仮説 H_0： $\mu=\mu_0$ を，尤度比検定によって，検定する。

尤度比

$$\lambda = \frac{\max_{\sigma^2} l(\mu_0,\sigma^2)}{\max_{\mu,\sigma^2} l(\mu,\sigma^2)} = \frac{l(\mu_0,\widetilde{\sigma}^2)}{l(\widehat{\mu},\widehat{\sigma}^2)}$$

について，制約の数は 1 なので，

$$-2\log\lambda \longrightarrow \chi^2(1)$$

となる。

ここで，$l(\mu,\sigma^2)$ は,

$$l(\mu,\sigma^2) = (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

であり，$\log l(\mu,\sigma^2)$ は，

$$\log l(\mu,\sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

である。

分子について：

$\mu=\mu_0$ のもとで，$\log l(\mu_0,\sigma^2)$ を σ^2 について最大化する。

$$\frac{\partial \log l(\mu_0,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu_0)^2 = 0$$

この解が $\widetilde{\sigma}^2$ である。したがって，

$$\widetilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2$$

を得る。よって，$l(\mu_0,\widetilde{\sigma}^2)$ は，

$$\begin{aligned}
&l(\mu_0,\widetilde{\sigma}^2)\\
&= (2\pi\widetilde{\sigma}^2)^{-n/2}\exp\left(-\frac{1}{2\widetilde{\sigma}^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right)\\
&= (2\pi\widetilde{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)
\end{aligned}$$

となる。

分母について：

問 (1) より，

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i$$
$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\widehat{\mu})^2$$

を得る。よって，$l(\widehat{\mu},\widehat{\sigma}^2)$ は，

$$\begin{aligned}
&l(\widehat{\mu},\widehat{\sigma}^2)\\
&= (2\pi\widehat{\sigma}^2)^{-n/2}\exp\left(-\frac{1}{2\widehat{\sigma}^2}\sum_{i=1}^{n}(x_i-\widehat{\mu})^2\right)\\
&= (2\pi\widehat{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)
\end{aligned}$$

となる。

したがって，

$$\begin{aligned}
\lambda &= \frac{\max_{\sigma^2} l(\mu_0,\sigma^2)}{\max_{\mu,\sigma^2} l(\mu,\sigma^2)}\\
&= \frac{l(\mu_0,\widetilde{\sigma}^2)}{l(\widehat{\mu},\widehat{\sigma}^2)}\\
&= \frac{(2\pi\widetilde{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)}{(2\pi\widehat{\sigma}^2)^{-n/2}\exp\left(-\frac{n}{2}\right)}\\
&= \left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\right)^{-n/2}
\end{aligned}$$

よって，近似的に，

$$-2\log\lambda = n(\log\widetilde{\sigma}^2 - \log\widehat{\sigma}^2) \sim \chi^2(1)$$

となる。

$-2\log\lambda > \chi^2_\alpha(1)$ のとき，有意水準 α で，帰無仮説 H_0： $\mu=\mu_0$ を棄却する。$\chi^2_\alpha(1)$ は自由度 1 のカイ二乗分布の $100\times\alpha$ % 点とする。

$\boxed{2}$ 次の問に答えよ。

(1) 離散型確率変数 X がベルヌイ分布に従うとき，その確率関数は次の式で表される。

$$f(x) = p^x(1-p)^{1-x} \qquad x=0,1$$

このベルヌイ分布から抽出された大きさ n の無作為標本を $X_1, X_2, \cdots, X_n$ とするとき，p の最尤推定量を求めなさい。

(2) Y を二項分布 $f(y)$ に従う確率変数とする。このとき，$\frac{Y}{n}$ は，n が大きくなると，p に近づくことを証明せよ。ただし，二項分布は

$$f(y) = {}_nC_y p^y(1-p)^{n-y} \qquad y=0,1,2,\cdots,n$$

で与えられる。

(3) 問 (2) の確率変数 Y について，確率変数 $Z_n \equiv \frac{Y-np}{\sqrt{np(1-p)}}$ を定義する。このとき，n が大きくなるにつれて，Z_n は標準正規分布に近づくことを証明せよ。

(4) 連続型確率変数 X の密度関数が

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \text{ のとき} \\ 0, & \text{その他} \end{cases}$$

とする。$\dfrac{X}{n}$ は，$n \longrightarrow \infty$ のとき，1 に近づくことを示せ。ただし，$\Gamma(a)$ はガンマ関数であり，

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} \mathrm{d}x$$

と定義される。

[**解答**]

(1) 離散型確率変数 X がベルヌイ分布に従うとき，その確率関数は次の式で表される。

$$f(x;p) = p^x (1-p)^{1-x} \qquad x = 0, 1$$

このベルヌイ分布から抽出された大きさ n の無作為標本を $X_1, X_2, \cdots, X_n$ とするとき，p の最尤推定量を求める。

$$\begin{aligned} & f(x_1, x_2, \cdots, x_n; p) \\ &= \prod_{i=1}^n f(x_i; p) \\ &= \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} \\ &= p^{\sum_i x_i} (1-p)^{n - \sum_i x_i} \\ &= l(p) \end{aligned}$$

対数をとる。

$$\begin{aligned} & \log l(p) \\ &= (\sum_i x_i) \log(p) + (n - \sum_i x_i) \log(1-p) \end{aligned}$$

対数尤度関数 $\log l(p)$ を p について微分して，ゼロと置く。

$$\begin{aligned} \frac{\mathrm{d} \log l(p)}{\mathrm{d}p} &= \frac{\sum_i x_i}{p} - \frac{n - \sum_i x_i}{1-p} \\ &= \frac{\sum_i x_i - np}{p(1-p)} \\ &= 0 \end{aligned}$$

この方程式を解く。

$$p = \frac{1}{n} \sum_{i=1}^n x_i = \overline{x}$$

p の最尤推定量は，

$$\widehat{p} = \frac{1}{n} \sum_{i=1}^n X_i = \overline{X}$$

となる。

(2) Y を二項分布 $f(y)$ に従う確率変数とするとき，$\dfrac{Y}{n}$ は，n が大きくなると，p に近づくことを証明する。

Y の平均，分散は，

$$\mathrm{E}(Y) = np, \qquad \mathrm{V}(Y) = np(1-p)$$

なので，

$$\begin{aligned} \mathrm{E}(\frac{Y}{n}) &= \frac{1}{n}\mathrm{E}(Y) = p \\ \mathrm{V}(\frac{Y}{n}) &= \frac{1}{n^2}\mathrm{V}(Y) = \frac{p(1-p)}{n} \end{aligned}$$

となる。

チェビシェフの不等式：

確率変数 X と $g(x) \geq 0$ について，

$$P\big(g(X) \geq k\big) \leq \frac{\mathrm{E}\big(g(X)\big)}{k}$$

となる。$k > 0$ とする。

ここで，$g(X) = \big(X - \mathrm{E}(X)\big)^2$, $k = \epsilon^2$ とすると，

$$P\big(|X - \mathrm{E}(X)| \geq \epsilon\big) \leq \frac{\mathrm{V}(X)}{\epsilon^2}$$

を得る。$\epsilon > 0$ とする。

X を $\dfrac{Y}{n}$ に置き換えて，そのままチェビシェフの不等式を当てはめると，

$$P\big(|\frac{Y}{n} - \mathrm{E}(\frac{Y}{n})| \geq \epsilon\big) \leq \frac{\mathrm{V}(\frac{Y}{n})}{\epsilon^2}$$

すなわち，$n \longrightarrow \infty$ のとき，

$$P\big(|\frac{Y}{n} - p| \geq \epsilon\big) \leq \frac{p(1-p)}{n\epsilon^2} \;\longrightarrow\; 0$$

となる。したがって，

$$\frac{Y}{n} \longrightarrow p$$

を得る。

(3) n 個の無作為標本 $X_1,\ X_2,\ \cdots,\ X_n$ はそれぞれ独立に，同一のベルヌイ分布に従うものとする。ただし，$P(X_i = x) = p^x(1-p)^{1-x},\ x = 0, 1$ とする。

$Y = X_1 + X_2 + \cdots + X_n$ を定義すると，Y は二項分布に従うので，$\dfrac{Y}{n}$ は標本平均とみなすことが出来る。

すなわち，$\dfrac{Y}{n} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$

よって，$\mathrm{E}(\dfrac{Y}{n}) = p,\ \mathrm{V}(\dfrac{Y}{n}) = p(1-p)/n$ を用いて，中心極限定理により，$n \longrightarrow \infty$ のとき，

$$\frac{\frac{Y}{n} - p}{\sqrt{p(1-p)/n}} \longrightarrow N(0,1)$$

を得る。

$$Z_n \equiv \frac{Y - np}{\sqrt{np(1-p)}} = \frac{\frac{Y}{n} - p}{\sqrt{p(1-p)/n}}$$

なので，

$$Z_n \longrightarrow N(0,1)$$

を得る。

(4) $X \sim \chi^2(n)$ のとき，

$\mathrm{E}(X) = n,\ \mathrm{V}(X) = 2n$ となるので，

$\mathrm{E}(\dfrac{X}{n}) = 1,\ \mathrm{V}(\dfrac{X}{n}) = \dfrac{2}{n}$ となる。

チェビシェフの不等式を当てはめる。

$$P\big(|\frac{X}{n} - \mathrm{E}(\frac{X}{n})| \geq \epsilon\big) \leq \frac{\mathrm{V}(\frac{X}{n})}{\epsilon^2}$$

$\epsilon > 0$ とする。すなわち，$n \longrightarrow \infty$ のとき，

$$P\big(|\frac{X}{n} - 1| \geq \epsilon\big) \leq \frac{2}{n\epsilon^2} \longrightarrow 0$$

となる。したがって，

$$\frac{X}{n} \longrightarrow 1$$

を得る。

$\boxed{3}$ 指数分布から生成された n 個の互いに独立な確率変数 $X_1,\ X_2,\ \cdots,\ X_n$ を考える。ただし，指数分布とは次の分布である。

$$f(x) = \lambda e^{-\lambda x} \qquad x > 0$$

で与えられる。このとき，次の問に答えよ。

(1) λ 最尤推定量を $\widehat{\lambda}$ とするとき，$\widehat{\lambda}$ を求めよ。

(2) n が大きいとき，$\widehat{\lambda}$ の平均，分散を求めよ。

[解答]

(1) $X_1, \cdots, X_n$ は互いに独立で、指数分布 $f(x)$ に従うので，尤度関数 $l(\lambda)$ は

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{n} f(x_i) \\ &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \\ &= \lambda^n e^{-\lambda \sum x_i} \end{aligned}$$

対数尤度関数は

$$\log l(\lambda) = n \log(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

$\log l(\lambda)$ を最大にするような λ を求める。

$$\frac{\mathrm{d} \log l(\lambda)}{\mathrm{d}\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

を解いて，x_i を X_i で置き換えて，λ の最尤推定量 $\widehat{\lambda}$ は

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n} X_i}$$

となる。

(2) n 個の無作為標本 $X_1,\ X_2,\ \cdots,\ X_n$

X_i の密度関数 $f(X_i; \lambda)$

母数 λ の最尤推定量 $\widehat{\lambda}_n$ について，$n \longrightarrow \infty$ のとき，

$$\sqrt{n}(\widehat{\lambda}_n - \lambda) \longrightarrow N\Big(0, \sigma^2\Big)$$

となる。ただし，

$$\sigma^2 = \sigma^2(\lambda) = \frac{1}{\mathrm{E}\left[\left(\dfrac{\mathrm{d}\log f(X;\lambda)}{\mathrm{d}\lambda}\right)^2\right]}$$

とする。

まず，$\sigma^2 = \sigma^2(\lambda)$ を求める。

$$\begin{aligned}
&\mathrm{E}\left[\left(\frac{\mathrm{d}\log f(X;\lambda)}{\mathrm{d}\lambda}\right)^2\right] \\
&= \mathrm{E}\left[\left(\frac{1}{\lambda} - X\right)^2\right] \\
&= \mathrm{E}\left(\frac{1}{\lambda^2} - \frac{2}{\lambda}X + X^2\right) \\
&= \frac{1}{\lambda^2} - \frac{2}{\lambda}\mathrm{E}(X) + \mathrm{E}(X^2) \\
&= \frac{1}{\lambda^2}
\end{aligned}$$

$\mathrm{E}(X)$, $\mathrm{E}(X^2)$ は

$$\mathrm{E}(X) = \frac{1}{\lambda}, \qquad \mathrm{E}(X^2) = \frac{2}{\lambda^2}$$

となる。

したがって，

$$\sigma^2 = \sigma^2(\lambda) = \frac{1}{\mathrm{E}\left[\left(\dfrac{\mathrm{d}\log f(X;\lambda)}{\mathrm{d}\lambda}\right)^2\right]} = \lambda^2$$

よって，n が大きいとき，

$$\widehat{\lambda}_n \sim N\left(\lambda, \frac{\lambda^2}{n}\right)$$

と近似して用いる。

$\boxed{4}$ $X_1, X_2, \cdots, X_n$ の n 個の確率変数は，それぞれ独立に，平均 μ，分散 σ^2 の分布をするものとする。次のような 2 つの μ の推定量 $\overline{X}, \widetilde{X}$ を考える。

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \widetilde{X} = \frac{1}{2}(X_1 + X_n)$$

このとき，次の問に答えよ。

(1) $\overline{X}, \widetilde{X}$ は不偏性を持つかどうかを調べよ。

(2) $\overline{X}, \widetilde{X}$ のどちらが有効かを調べよ。

(3) $\overline{X}, \widetilde{X}$ の一致性について調べよ。

[解答]

(1) $\overline{X}, \widetilde{X}$ は不偏性を持つかどうかを調べる。

$$\begin{aligned}
\mathrm{E}(\overline{X}) &= \mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) \\
&= \frac{1}{n}\mathrm{E}\Big(\sum_{i=1}^{n} X_i\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mu \\
&= \mu
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}(\widetilde{X}) &= \frac{1}{2}\Big(\mathrm{E}(X_1) + \mathrm{E}(X_n)\Big) \\
&= \frac{1}{2}(\mu + \mu) \\
&= \mu
\end{aligned}$$

よって，両方とも μ の不偏推定量である。

(2) $\overline{X}, \widetilde{X}$ のどちらが有効かを調べる。

$$\begin{aligned}
\mathrm{V}(\overline{X}) &= \mathrm{V}\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) \\
&= \frac{1}{n^2}\mathrm{V}\Big(\sum_{i=1}^{n} X_i\Big) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}(X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

$$\begin{aligned}
\mathrm{V}(\widetilde{X}) &= \frac{1}{4}\Big(\mathrm{V}(X_1) + \mathrm{V}(X_n)\Big) \\
&= \frac{1}{4}(\sigma^2 + \sigma^2) \\
&= \frac{\sigma^2}{2}
\end{aligned}$$

なので，$n > 2$ のとき，$\mathrm{V}(\overline{X}) < \mathrm{V}(\widetilde{X})$ となり，$\overline{X}$ が $\widetilde{X}$ より有効となる。

(3) $\overline{X}, \widetilde{X}$ の一致性について調べる。

チェビシェフの不等式を当てはめる。$\overline{X}$ について，

$$P\big(|\overline{X}-\mathrm{E}(\overline{X})| \geq \epsilon\big) \leq \frac{\mathrm{V}(\overline{X})}{\epsilon^2}$$

$\epsilon > 0$ とする。すなわち，$n \longrightarrow \infty$ のとき，

$$P\big(|\overline{X}-\mu| \geq \epsilon\big) \leq \frac{\sigma^2}{n\epsilon^2} \longrightarrow 0$$

となる。したがって，

$$\overline{X} \longrightarrow \mu$$

を得る。

一方，$\widetilde{X}$ について，

$$P\big(|\widetilde{X}-\mathrm{E}(\widetilde{X})| \geq \epsilon\big) \leq \frac{\mathrm{V}(\widetilde{X})}{\epsilon^2}$$

$\epsilon > 0$ とする。すなわち，$n \longrightarrow \infty$ のとき，

$$P\big(|\widetilde{X}-\mu| \geq \epsilon\big) \leq \frac{\sigma^2}{2\epsilon^2} > 0$$

となる。

$\overline{X}$ は μ の一致推定量であるが，$\widetilde{X}$ は一致性はない。

$\boxed{5}$ 正規母集団 $N(\mu, \sigma^2)$ から 9 個の無作為標本

21 23 32 20 36 27 26 28 30

が得らた。このとき、次の各問に答えなさい。

(1) μ と σ^2 の不偏推定値を求めよ。

(2) 信頼係数 0.90 および 0.95 の μ の信頼区間を求めよ。

(3) 帰無仮説 H_0：$\mu = 24$ を対立仮説 H_1：$\mu > 24$ に対して有意水準 0.10 および 0.05 で検定しなさい。

[解答]

(1) μ と σ^2 の不偏推定量 $\overline{X}, S^2$ は

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

で，不偏推定値は

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

である。

$$\begin{aligned}\overline{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ &= \frac{1}{9}(21+23+32+20 \\ &\quad +36+27+26+28+30) \\ &= 27\end{aligned}$$

$$\begin{aligned}s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \\ &= \frac{1}{8}\Big((21-27)^2+(23-27)^2+(32-27)^2 \\ &\quad +(20-27)^2+(36-27)^2+(27-27)^2 \\ &\quad +(26-27)^2+(28-27)^2+(30-27)^2\Big) \\ &= \frac{1}{8}(36+16+25+49 \\ &\quad +81+0+1+1+9) \\ &= 27.25\end{aligned}$$

(2) μ の信頼区間を求める。

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1),$$

を利用する。

$$P\left(\left|\frac{\overline{X}-\mu}{S/\sqrt{n}}\right| < t_{\alpha/2}(n-1)\right) = 1-\alpha$$

$t_{\alpha/2}(n-1)$ は $100 \times \dfrac{\alpha}{2}$ % 点で，確率 α と自由度 $n-1$ が与えられると，t 分布表から得られる。

したがって，

$$P\left(\overline{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$
$$= 1-\alpha$$

$\overline{X}, S^2$ を $\overline{x}, s^2$ で置き換えて，

信頼係数 $1-\alpha$ の μ の信頼区間：$\Longrightarrow$

$$\left(\overline{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \overline{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right)$$

さらに，$\overline{x} = 27$, $s^2 = 27.25$, $n = 9$, $t_{0.05}(8) = 1.860$, $t_{0.025}(8) = 2.306$ なので，

信頼係数 0.90 の μ の信頼区間は，

$(27-1.860\sqrt{\dfrac{27.25}{9}}, 27+1.860\sqrt{\dfrac{27.25}{9}})$
$=(23.76, 30.24)$

となり，信頼係数 0.95 の μ の信頼区間は,

$(27-2.306\sqrt{\dfrac{27.25}{9}}, 27+2.306\sqrt{\dfrac{27.25}{9}})$
$=(22.99, 31.01)$

となる。

(3) 帰無仮説 H_0 : $\mu=24$ を対立仮説 H_1 : $\mu>24$ に対して有意水準 0.10 および 0.05 で検定する。

$\overline{X}$ の分布は，$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ なので，帰無仮説 H_0 : $\mu=\mu_0$ が正しいもとで，$\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$ となる (μ を μ_0 で置き換える)。このとき，検定統計量 $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$。検定統計量の値 $\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}$。

対立仮説 H_1 : $\mu>\mu_0$ (片側検定) について：

$P\Big(\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} > t_\alpha(n-1)\Big)=\alpha$ なので,

$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} > t_\alpha(n-1)$ のとき,

有意水準 α で H_0 : $\mu=\mu_0$ を棄却する。

$\overline{x}=27$, $s^2=27.25$, $\mu_0=24$, $n=9$, $t_{0.10}(8)=1.397$, $t_{0.05}(8)=1.860$ を当てはめると,

$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} = \dfrac{27-24}{\sqrt{27.25/9}} = 1.724 > t_{0.10}(8)=1.397$ となり，有意水準 0.10 で H_0 : $\mu=24$ を棄却する。

$\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} = \dfrac{27-24}{\sqrt{27.25/9}} = 1.724 < t_{0.05}(8)=1.860$ となり，有意水準 0.05 で H_0 : $\mu=24$ を棄却しない。

$\boxed{6}$ 平均 μ，分散が既知で $\sigma^2=2^2$ である正規母集団から 16 個の無作為標本 $X_1, X_2, \cdots, X_{16}$ を抽出し，その標本平均を計算したところ，$\overline{x}=36$ であった。このとき，次の各問に答えよ。

(1) 平均 μ の信頼係数 0.95 の信頼区間を求めよ。

(2) 帰無仮説 H_0 : $\mu=35$ を対立仮説 H_1 : $\mu=36.5$ に対して，有意水準 0.05 で検定せよ。

(3) 問 (2) で行った検定の検出力を求めよ。

[解答]

(1) 平均 μ の信頼係数 0.95 の信頼区間を求める。

$\overline{X}$ の分布は

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

となる。

$$P\Big(\Big|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\Big| < z_{\alpha/2}\Big)=1-\alpha$$

$z_{\alpha/2}$ は $100\times\dfrac{\alpha}{2}$ % 点で，確率 α が与えられると，正規分布表から得られる。

したがって,

$P\Big(\overline{X}-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} < \mu < \overline{X}+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\Big)=1-\alpha$

$\overline{X}$ を $\overline{x}$ で置き換えて,

信頼係数 $1-\alpha$ の μ の信頼区間：$\Longrightarrow$
$\Big(\overline{x}-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \overline{x}+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\Big)$

$\overline{x}=36$, $\sigma^2=2^2$, $n=16$, $z_{0.025}=1.960$ を代入すると,

平均 μ の信頼係数 0.95 の信頼区間は,

$(36-1.960\dfrac{2}{\sqrt{16}}, 36+1.960\dfrac{2}{\sqrt{16}})=(35.02, 36.98)$

となる。

(2) 帰無仮説 H_0 : $\mu=35$ を対立仮説 H_1 : $\mu=36.5$ に対して，有意水準 0.05 で検定する。

$\overline{X}$ の分布は，$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ なので，帰無仮説 H_0 : $\mu=\mu_0$ が正しいもとで,

$\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ となる (μ を μ_0 で置き換える)。このとき，検定統計量 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$。検定統計量の値 $\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$。

対立仮説 H_1 : $\mu>\mu_0$ (片側検定) について：

$P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha\Big)=\alpha$ なので，$\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha$ のとき，有意水準 α で帰無仮説 H_0 : $\mu=\mu_0$ を棄却する。

$\overline{x}=36$, $\sigma^2=2^2$, $n=16$, $z_{0.05}=1.645$ を代入すると,
$\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{36-35}{2/\sqrt{16}} = 2 > z_\alpha=1.645$ なので，有意水準 $\alpha=0.05$ で帰無仮説 H_0 : $\mu=35$ を棄却する。

(3) 問 (2) で行った検定の検出力を求める。

検出力とは，対立仮説のもとで，帰無仮説を棄却する確率である。

すなわち，帰無仮説 H_0： $\mu=\mu_0$ のもとで，帰無仮説を棄却する棄却域は，$P\Big(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}>z_\alpha\Big)=\alpha$ なので，

$\overline{X}>\mu_0+z_\alpha\sigma/\sqrt{n}$ となる。

対立仮説 H_1： $\mu=\mu_1$ のもとで，帰無仮説を棄却する確率を求める。

すなわち，対立仮説 H_1： $\mu=\mu_1$ のもとで，
$P\Big(\overline{X}>\mu_0+z_\alpha\sigma/\sqrt{n}\Big)$ を求める。
対立仮説 H_1： $\mu=\mu_1$ のもとで，
$\dfrac{\overline{X}-\mu_1}{\sigma/\sqrt{n}}\sim N(0,1)$ となるので，
$P\Big(\dfrac{\overline{X}-\mu_1}{\sigma/\sqrt{n}}>\dfrac{\mu_0-\mu_1}{\sigma/\sqrt{n}}+z_\alpha\Big)$
となる確率を求めればよい。

$\sigma=2$, $n=16$, $\mu_0=35$, $\mu_1=36.5$, $z_\alpha=1.645$ を代入すると，

$$
\begin{aligned}
&P\Big(\frac{\overline{X}-\mu_1}{\sigma/\sqrt{n}}>\frac{35-36.5}{2/\sqrt{16}}+1.645\Big)\\
&=P\Big(\frac{\overline{X}-\mu_1}{\sigma/\sqrt{n}}>-1.355\Big)\\
&=1-0.0877=0.9123
\end{aligned}
$$

を得る。

($z_{0.0885}=1.35$, $z_{0.0869}=1.36$ に注意)

$\boxed{7}$ $X_1, X_2, \cdots, X_n$ は互いに独立で，すべて同一のポアソン分布に従うものとする。ただし，ポアソン分布の確率関数は

$$P(X=x)=f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!},\quad x=0,1,2,\cdots$$

である。
このとき，次の問に答えよ。

(1) λ の最尤推定量 $\widehat{\lambda}$ を求めよ。
(2) $\widehat{\lambda}$ は，λ の不偏推定量であることを証明せよ。
(3) $\widehat{\lambda}$ は，λ の有効推定量であることを証明せよ。
(4) $\widehat{\lambda}$ は，λ の一致推定量であることを証明せよ。

[解答]

(1) λ の最尤推定量 $\widehat{\lambda}$ を求める。

ポアソン分布の確率関数は,

$$P(X=x)=f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!},\quad x=0,1,2,\cdots$$

なので，尤度関数は,

$$
\begin{aligned}
l(\lambda)&=\prod_{i=1}^n f(x_i;\lambda)\\
&=\prod_{i=1}^n \frac{\lambda^{x_i}e^{-\lambda}}{x_i!}\\
&=\frac{\lambda^{\sum_{i=1}^n x_i}e^{-n\lambda}}{\prod_{i=1}^n x_i!}
\end{aligned}
$$

対数尤度関数は,

$$\log l(\lambda)=\log(\lambda)\sum_{i=1}^n x_i-n\lambda-\log(\prod_{i=1}^n x_i!)$$

となる。

$$
\begin{aligned}
\frac{\partial\log l(\lambda)}{\partial\lambda}&=\frac{1}{\lambda}\sum_{i=1}^n x_i-n\\
&=0
\end{aligned}
$$

これを解いて，λ の最尤推定量 $\widehat{\lambda}$ は,

$$\widehat{\lambda}=\frac{1}{n}\sum_{i=1}^n X_i=\overline{X}$$

となる。

(2) $\widehat{\lambda}$ は，λ の不偏推定量であることを証明する。

$$
\begin{aligned}
\mathrm{E}(\widehat{\lambda})&=\mathrm{E}(\frac{1}{n}\sum_{i=1}^n X_i)\\
&=\frac{1}{n}\sum_{i=1}^n \mathrm{E}(X_i)\\
&=\frac{1}{n}\sum_{i=1}^n \lambda\\
&=\lambda
\end{aligned}
$$

(3) $\widehat{\lambda}$ は，λ の有効推定量であることを証明する。

クラメール・ラオの不等式の等号が成り立つことを証明すればよい。

$$\begin{aligned}
\mathrm{V}(\widehat{\lambda}) &= \mathrm{V}(\frac{1}{n}\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{V}(X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \lambda \\
&= \frac{\lambda}{n}
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\lambda)}{\partial\lambda}\right)^2\right]} \\
&= \frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial (X\log\lambda - \lambda - \log X!)}{\partial\lambda}\right)^2\right]} \\
&= \frac{1}{n\mathrm{E}\left[\left(\dfrac{X}{\lambda} - 1\right)^2\right]} \\
&= \frac{\lambda^2}{n\mathrm{E}[(X-\lambda)^2]} \\
&= \frac{\lambda^2}{n\mathrm{V}(X)} \\
&= \frac{\lambda^2}{n\lambda} \\
&= \frac{\lambda}{n}
\end{aligned}$$

したがって，

$$\mathrm{V}(\widehat{\lambda}) = \frac{1}{n\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\lambda)}{\partial\lambda}\right)^2\right]}$$

となり，$\mathrm{V}(\widehat{\lambda})$ は，クラメール・ラオの下限に一致する。よって，$\widehat{\lambda}$ は有効推定量である。

(4) $\widehat{\lambda}$ は，λ の一致推定量であることを証明する。

$$\mathrm{E}(\widehat{\lambda}) = \lambda, \qquad \mathrm{V}(\widehat{\lambda}) = \frac{\lambda}{n}$$

である。チェビシェフの不等式

$$P\big(|\widehat{\lambda} - \mathrm{E}(\widehat{\lambda})| \geq \epsilon\big) \leq \frac{\mathrm{V}(\widehat{\lambda})}{\epsilon^2}$$

に，$\mathrm{E}(\widehat{\lambda})$, $\mathrm{V}(\widehat{\lambda})$ を代入すると，

$$P(|\widehat{\lambda} - \lambda| > \epsilon) < \frac{\lambda}{n\epsilon^2} \;\longrightarrow\; 0$$

が得られる。したがって，一致性も成り立つ。

$\boxed{8}$ 連続型確率変数 $X_1, X_2, \cdots, X_n$ は互いに独立に同一の正規分布に従うものとする。このとき，以下の問に答えよ。ただし，正規分布の密度関数は

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

で表される。

(1) 標本平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ の分布は，平均 μ，分散 $\dfrac{\sigma^2}{n}$ の正規分布に従うことを示せ。

(2) $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ を定義する。Z の分布は，平均 0，分散 1 の正規分布に従うことを示せ。

(3) 標本不偏分散 $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$ を考える。$\dfrac{(n-1)S^2}{\sigma^2}$ の分布は自由度 $n-1$ のカイ二乗分布であることが知られている。これを利用して，S^2 の平均と分散を求めよ。

ただし，自由度 m のカイ自乗分布の密度関数は，

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}}, & x > 0 \text{ のとき} \\ 0, & \text{その他} \end{cases}$$

として表される。

(4) S^2 は σ^2 の一致推定量であることを示せ。

[解答]

(1) 標本平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ の分布は，平均 μ，分散 $\dfrac{\sigma^2}{n}$ の正規分布に従うことを示す。

積率母関数を利用する。

$X \sim N(\mu, \sigma^2)$ のとき，X の積率母関数 $\phi(\theta)$ は，

$$
\begin{aligned}
&\phi(\theta)\\
&\equiv \mathrm{E}(e^{\theta X})\\
&= \int_{-\infty}^{\infty} e^{\theta x} f(x) \mathrm{d}x\\
&= \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \mathrm{d}x\\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2+\theta x} \mathrm{d}x\\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x^2-2(\mu+\sigma^2\theta)x+\mu^2\right)} \mathrm{d}x\\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x-(\mu+\sigma^2\theta)\right)^2+(\mu\theta+\frac{1}{2}\sigma^2\theta^2)} \mathrm{d}x\\
&= e^{\mu\theta+\frac{1}{2}\sigma^2\theta^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(x-(\mu+\sigma^2\theta)\right)^2} \mathrm{d}x\\
&= \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right)
\end{aligned}
$$

と計算される。積分のところは，$N(\mu+\sigma^2\theta, \sigma^2)$ の確率密度関数に注意

よって，$X_i \sim N(\mu, \sigma^2)$ のとき，X_i の積率母関数 $\phi_i(\theta)$ は，

$$
\phi_i(\theta) = \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right)
$$

となる。

今，$\overline{X}$ の積率母関数 $\phi_{\overline{x}}(\theta)$ を考える。

$$
\begin{aligned}
\phi_{\overline{x}}(\theta) &\equiv \mathrm{E}(e^{\theta\overline{X}})\\
&= \mathrm{E}(e^{\theta\frac{1}{n}\sum_{i=1}^{n} X_i})\\
&= \mathrm{E}(\prod_{i=1}^{n} e^{\frac{\theta}{n}X_i})\\
&= \prod_{i=1}^{n} \mathrm{E}(e^{\frac{\theta}{n}X_i})\\
&= \prod_{i=1}^{n} \phi_i(\frac{\theta}{n})\\
&= \prod_{i=1}^{n} \exp\left(\mu\frac{\theta}{n} + \frac{1}{2}\sigma^2\left(\frac{\theta}{n}\right)^2\right)\\
&= \exp\left(\mu\theta + \frac{1}{2}\sigma^2\frac{\theta^2}{n}\right)\\
&= \exp\left(\mu\theta + \frac{1}{2}\frac{\sigma^2}{n}\theta^2\right)
\end{aligned}
$$

となり，これは，平均 μ，分散 σ^2/n の正規分布の積率母関数に一致する。

(2) $Z = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ の分布は，平均 0，分散 1 の正規分布に従うことを示す。

$\overline{X}$ の積率母関数 $\phi_{\overline{x}}(\theta)$ は，問 (1) より，

$$
\begin{aligned}
\phi_{\overline{x}}(\theta) &\equiv \mathrm{E}(e^{\theta\overline{X}})\\
&= \exp\left(\mu\theta + \frac{1}{2}\frac{\sigma^2}{n}\theta^2\right)
\end{aligned}
$$

と計算されることを利用する。

Z の積率母関数 $\phi_z(\theta)$ は，

$$
\begin{aligned}
\phi_z(\theta) &\equiv \mathrm{E}(e^{\theta Z})\\
&= \mathrm{E}\left(\exp(\theta\frac{\overline{X}-\mu}{\sigma/\sqrt{n}})\right)\\
&= \exp\left(-\theta\frac{\mu}{\sigma/\sqrt{n}}\right)\mathrm{E}\left(\exp(\frac{\theta}{\sigma/\sqrt{n}}\overline{X})\right)\\
&= \exp\left(-\theta\frac{\mu}{\sigma/\sqrt{n}}\right)\phi_{\overline{x}}(\frac{\theta}{\sigma/\sqrt{n}})\\
&= \exp\left(-\theta\frac{\mu}{\sigma/\sqrt{n}}\right)\\
&\quad \times \exp\left(\mu\frac{\theta}{\sigma/\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n}\left(\frac{\theta}{\sigma/\sqrt{n}}\right)^2\right)\\
&= \exp(\frac{1}{2}\theta^2)
\end{aligned}
$$

これは，$N(0,1)$ の積率母関数となっている。

(3) まず，準備として，自由度 m のカイ二乗分布の平均と分散を計算する。

自由度 m のカイ二乗分布は，

$$
f(x) = \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} \quad x > 0 \text{ のとき}
$$

なので，その積率母関数 $\phi_{\chi^2}(\theta)$ は，

$$
\begin{aligned}
&\phi_{\chi^2}(\theta)\\
&= \mathrm{E}(e^{\theta X})\\
&= \int_0^{\infty} e^{\theta x} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} \mathrm{d}x\\
&= \int_0^{\infty} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{1}{2}(1-2\theta)x} \mathrm{d}x\\
&= \int_0^{\infty} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} \left(\frac{y}{1-2\theta}\right)^{\frac{m}{2}-1} e^{-\frac{1}{2}y} \frac{1}{1-2\theta} \mathrm{d}x
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\frac{1}{1-2\theta}\right)^{\frac{m}{2}-1} \frac{1}{1-2\theta} \\
&\times \int_0^\infty \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m}{2}-1} e^{-\frac{1}{2}y} \mathrm{d}x \\
&= (1-2\theta)^{-\frac{m}{2}}
\end{aligned}
$$

となる。

4 つ目の等式で，$y = (1-2\theta)x$ として，置換積分を利用。

$\dfrac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m}{2}-1} e^{-\frac{1}{2}y}$ は，自由度 m の χ^2 分布となっているので，その積分値は 1 となる。

積率母関数を微分して，

$$
\begin{aligned}
\phi'_{\chi^2}(\theta) &= m(1-2\theta)^{-\frac{m}{2}-1} \\
\phi''_{\chi^2}(\theta) &= m(m+2)(1-2\theta)^{-\frac{m}{2}-2}
\end{aligned}
$$

を用いると，

$$
\begin{aligned}
\mathrm{E}(X) &= \phi'_{\chi^2}(0) = m \\
\mathrm{E}(X^2) &= \phi''_{\chi^2}(0) = m(m+2)
\end{aligned}
$$

が得られる。

よって，自由度 m の χ^2 分布の平均は m，分散は，

$$
\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 \\
&= m(m+2) - m^2 \\
&= 2m
\end{aligned}
$$

となる。

$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ なので，これを利用すると，

$$
\begin{aligned}
&\mathrm{E}(\frac{(n-1)S^2}{\sigma^2}) = n-1 \\
&\mathrm{V}(\frac{(n-1)S^2}{\sigma^2}) = 2(n-1)
\end{aligned}
$$

となる。よって，

$$
\begin{aligned}
&\frac{n-1}{\sigma^2}\mathrm{E}(S^2) = n-1 \\
&(\frac{n-1}{\sigma^2})^2\mathrm{V}(S^2) = 2(n-1)
\end{aligned}
$$

を利用して，S^2 の平均と分散は，

$$
\begin{aligned}
\mathrm{E}(S^2) &= \sigma^2 \\
\mathrm{V}(S^2) &= \frac{2\sigma^4}{n-1}
\end{aligned}
$$

となる。

(4) S^2 は σ^2 の一致推定量であることを示す。

チェビシェフの不等式を用いる。

$$
P\big(|S^2 - \mathrm{E}(S^2)| \geq \epsilon\big) \leq \frac{\mathrm{V}(S^2)}{\epsilon^2}
$$

に，$\mathrm{E}(S^2), \mathrm{V}(S^2)$ を代入すると，

$$
P\big(|S^2 - \sigma^2| \geq \epsilon\big) \leq \frac{2\sigma^4}{(n-1)\epsilon^2} \longrightarrow 0
$$

が得られる。したがって，一致性も成り立つ。