帰無仮説 H_0 が正しいもとで，

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0,1)$$

なので，

$$P\left(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < -z_\alpha\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$ を検定統計量と呼ぶ。

$\overline{X}_1$，$\overline{X}_2$ をその推定値（データ）$\overline{x}_1$，$\overline{x}_2$ で置き換えて，もし $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < -z_\alpha$ なら

ば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ **有意水準 α で，母平均 μ_1 は μ_2 よりも小さいと判断する。**

● ケース 2 (片側検定)

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 > \mu_2$

帰無仮説 H_0 が正しいもとで，

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$$

なので，

$$P\left(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > z_\alpha\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$ を検定統計量と呼ぶ。

$\overline{X}_1$，$\overline{X}_2$ をその推定値（データ）$\overline{x}_1$，$\overline{x}_2$ で置き換えて，もし $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > z_\alpha$ ならば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ **有意水準 α で，母平均 μ_1 は μ_2 よりも大きいと判断する。**

● **ケース 3 (両側検定)**

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 \neq \mu_2$

帰無仮説 H_0 が正しいもとで，

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$$

なので，

$$P\left(\left|\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right| > z_{\alpha/2}\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$ を検定統計量と呼ぶ。

$\overline{X}_1$，$\overline{X}_2$ をその推定値（データ）$\overline{x}_1$，$\overline{x}_2$ で置き換えて，もし

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < -z_{\alpha/2}$$

または

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > z_{\alpha/2}$$

ならば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母平均 μ_1 と μ_2 は異なると判断する。

例題 8.6 (P.147)：　あるデパートで，店員とアルバイトが同じ商品を包装し，一時間の作業で以下の結果を得た。

	人数	平均包装数
店員	5	64
アルバイト	9	56

店員の方がアルバイトより熟練していると言えるか？有意水準 5% で検定せよ。ただし，店員とアルバイトの包装数は $N(\mu_1, 30.5)$，$N(\mu_2, 75.6)$ に従うことが分かっているものとせよ。

解答：　添字 1 を店員，添字 2 をアルバイトとする。

$$\overline{X}_1 - \overline{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$$

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 > \mu_2$

帰無仮説 H_0 **が正しいもとで，**

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$$

なので，

$$P\Bigl(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > z_\alpha\Bigr) = \alpha$$

となる。

もし $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > z_\alpha$ ならば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母平均 μ_1 は μ_2 よりも大きいと判断する。

$\alpha = 0.05$ のとき，$z_\alpha = 1.645$ である。

また，$\overline{x}_1 = 64$，$\overline{x}_2 = 56$，$\sigma_1^2 = 30.5$，$\sigma_2^2 = 75.6$，$n_1 = 5$，$n_2 = 9$ なので，

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

$$= \frac{64 - 56}{\sqrt{30.5/5 + 75.6/9}}$$

$$= 2.101 > z_{0.05} = 1.645$$

となり，H_0 を棄却する。すなわち，店員の方がアルバイトより熟練しているといえる。

8.6.2 母分散が未知の場合 (分布を仮定しない，n_1, n_2 共に大きいとき, P.148 の真中)

2 つのグループ

・第 1 グループ：

大きさ n_1 の無作為標本 $X_{11}, X_{12}, \cdots, X_{1n_1}$

$\mathbf{E}(X_{1i}) = \mu_1$，$\mathbf{V}(X_{1i}) = \sigma_1^2$，$i = 1, 2, \cdots, n_1$

標本平均 $\overline{X}_1$

標本不偏分散 $S_1^2 = \dfrac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \overline{X}_1)^2$

・第 2 グループ：

大きさ n_2 **の無作為標本** $X_{21}, X_{22}, \cdots, X_{2n_2}$

$\mathbf{E}(X_{2i}) = \mu_2$，$\mathbf{V}(X_{2i}) = \sigma_2^2$，　$i = 1, 2, \cdots, n_2$

標本平均 $\overline{X}_2$

標本不偏分散 $S_2^2 = \dfrac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \overline{X}_2)^2$

n_1, n_2 **共に大きいとき, 母分散が未知でも既知でも，正規分布の仮定は必要ない。**

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。**

母分散が既知の場合は，中心極限定理にもとづいて，

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$$

を使う。

母分散が未知の場合も，σ_1^2，σ_2^2 を S_1^2，S_2^2 で置き換えて，n_1, n_2 が大きいとき中心極限定理 (定理 6.1, P.90) によって，近似的に

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0, 1)$$

となる。ただし，

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \overline{X}_1)^2$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \overline{X}_2)^2$$

とする。

注意) この場合，t 分布にはならない。

理由)

t 分布： $Z \sim N(0, 1)$，$U \sim \chi^2(k)$，Z と U は独立のとき，$T = \dfrac{Z}{\sqrt{U/k}} \sim t(k)$ となる。

この場合，$Z = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$

(*) $U \sim \chi^2(m)$，$V \sim \chi^2(n)$，U と V は独立のとき，$U + V \sim \chi^2(m + n)$ となる。

$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1)$，$\frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_1-1)$，$S_1^2$ と S_2^2 は異なるグループで独立なので，

$U = \frac{(n_1-1)S_1^2}{\sigma_1^2} + \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_1+n_2-2)$ となる。

$$T = \frac{Z}{\sqrt{U/k}} = \frac{\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}}{\sqrt{\left(\dfrac{(n_1-1)S_1^2}{\sigma_1^2}+\dfrac{(n_2-1)S_2^2}{\sigma_2^2}\right)\Big/(n_1+n_2-2)}} \sim t(n_1+n_2-2)$$ となるが，σ_1^2，σ_2^2

が残るため，区間推定や仮説検定ができない。

例外)

$\sigma_1^2=\sigma_2^2=\sigma^2$ を仮定するときだけ，$\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}} \sim t(n_1+n_2-2)$ と

なり，σ^2 が消えてなくなるので，$\mu_1 - \mu_2$ に関する t 分布による通常の区間推定，仮説検定が行えるようになる。

● ケース 1 (片側検定)

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 < \mu_2$

帰無仮説 H_0 が正しいもとで，近似的に，

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0, 1)$$

なので，

$$P\left(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} < -z_\alpha\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ を検定統計量と呼ぶ。

もし $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} < -z_\alpha$ ならば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母平均 μ_1 は μ_2 よりも小さいと判断する。

ただし，

$$s_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2$$

$$s_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(x_{2i} - \overline{x}_2)^2$$

とする。

● ケース 2 (片側検定)

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 > \mu_2$

帰無仮説 H_0 が正しいもとで，近似的に，

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0,1)$$

なので，

$$P\Bigl(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} > z_\alpha\Bigr) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ を検定統計量と呼ぶ。

もし $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} > z_\alpha$ ならば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということに

なり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母平均 μ_1 は μ_2 よりも大きいと判断する。

● ケース 3 (両側検定)

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 \neq \mu_2$

帰無仮説 H_0 が正しいもとで，近似的に，

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0, 1)$$

なので，

$$P\left(\left|\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}\right| > z_{\alpha/2}\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ を検定統計量と呼ぶ。

もし

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} < -z_{\alpha/2}$$

または

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} > z_{\alpha/2}$$

ならば，帰無仮説 $H_0 : \mu_1 = \mu_2$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母平均 μ_1 と μ_2 は異なると判断する。

例題 8.7 (P.148)：　A地方とB地方とで，一年間の収入に差があるかどうかを調べたい。

	標本数	標本平均	標準偏差
A	154	615	40
B	120	606	32

有意水準 5 % で，A 地方の年間収入が B 地方と異なるかどうかを検定せよ。

解答：　添字 1 を A 地方，添字 2 を B 地方とする。

n_1, n_2 が大きいとき，中心極限定理で近似的に，

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0, 1)$$

となる。

帰無仮説　$H_0 : \mu_1 = \mu_2$

対立仮説　$H_1 : \mu_1 \neq \mu_2$

帰無仮説 H_0 **が正しいもとで，** n_1, n_2 **が大きいとき，近似的に，**

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0, 1)$$

なので，

$$P\left(\left|\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}\right| > z_{\alpha/2}\right) = \alpha$$

となる。

このとき， $\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ **を検定統計量と呼ぶ。**

もし $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} < -z_{\alpha/2}$**，または，**$\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} > z_{\alpha/2}$ **ならば，帰無仮説** $H_0 : \mu_1 = \mu_2$

が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母平均 μ_1 は μ_2 と異なると判断する。

$\alpha = 0.05$ のとき，$z_{\alpha/2} = 1.960$ である。

また，$\overline{x}_1 = 615$，$\overline{x}_2 = 606$，$s_1^2 = 40^2$，$s_2^2 = 32^2$，$n_1 = 154$，$n_2 = 120$ なので，

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

$$= \frac{615 - 606}{\sqrt{40^2/154 + 32^2/120}}$$

$$= 2.07 > z_{0.025} = 1.960$$

となり，有意水準 0.05 で，H_0 を棄却する。

すなわち，A地方とB地方の平均収入には差があるといえる。

8.7　等分散の検定 (P.149) *** 時間がなければ省略

P.98の「6.4.3　F 分布」に戻る。

定理 6.6

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

このとき，

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

自由度 (m,n) の F 分布

定理 6.7

2 つの正規母集団 $N(\mu_1,\sigma^2)$，$N(\mu_2,\sigma^2)$ から無作為標本（第 1 グループ，第 2 グループ）

標本の大きさ n_1，n_2

標本不偏分散 $S_1^2 = \dfrac{1}{n_1-1}\displaystyle\sum_{i=1}^{n_1}(X_{1i}-\overline{X}_1)^2$，$S_2^2 = \dfrac{1}{n_2-1}\displaystyle\sum_{j=1}^{n_2}(X_{2j}-\overline{X}_2)^2$

第 1 グループの無作為標本： $\dfrac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2(n_1-1)$

第 2 グループの無作為標本： $\dfrac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_2-1)$

このとき，定理 6.6 から，

$$\frac{\dfrac{(n_1-1)S_1^2}{\sigma^2}\Big/(n_1-1)}{\dfrac{(n_2-1)S_2^2}{\sigma^2}\Big/(n_2-1)} = \frac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1)$$

となる。

F 分布の表 P.254-255

横は分子の自由度，縦は分母の自由度

P.254 は上側 5 % 点，P.255 は上側 1 % 点

F 分布の性質：

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

$Y = \dfrac{U/m}{V/n} \sim F(m,n)$ となる（定理 6.6）。

$$\Longrightarrow \quad \frac{1}{Y} = \frac{V/n}{U/m} \sim F(n,m)$$

$\Longrightarrow$ すなわち，逆数の分布は，自由度を入れ替えて F 分布

$F_\alpha(m,n)$ を自由度 m, n の F 分布の上側 $100 \times \alpha$ % 点とする。

すなわち，$Y \sim F(m,n)$，上側確率 α のとき，$F_\alpha(m,n)$ は，

$$P\Big(Y > F_\alpha(m,n)\Big) = \alpha \qquad \Longrightarrow \qquad P\Big(Y \leq F_\alpha(m,n)\Big) = P\Big(Y < F_\alpha(m,n)\Big) = 1 - \alpha$$

となる数値である（$\alpha = 0.05,\ 0.01$ のとき，付表 4 から得られる）。

また，

$$P\Big(\frac{1}{Y} > F_{\alpha}(n,m)\Big) = \alpha \qquad \Longrightarrow \qquad P\Big(Y < \frac{1}{F_{\alpha}(n,m)}\Big) = \alpha$$

となる。

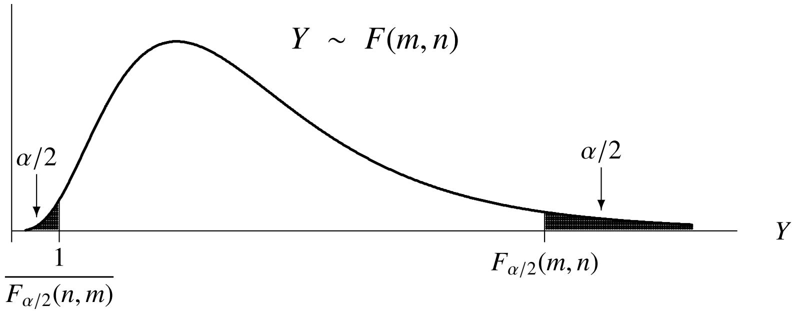
すなわち，

$$P\Big(Y > F_{\alpha/2}(m,n)\Big) = \alpha/2 \qquad \Longrightarrow \qquad P\Big(Y < \frac{1}{F_{\alpha/2}(n,m)}\Big) = \alpha/2$$

となる。したがって，

$$P\Big(\frac{1}{F_{\alpha/2}(n,m)} < Y < F_{\alpha/2}(m,n)\Big) = 1 - \alpha$$

が得られる（分散に関する仮説検定で利用）。



P.149 に戻る。

● 分散が等しいかどうかの検定

2 つの標本は正規母集団からの無作為標本

第 1 グループ： 大きさ n_1 **の無作為標本，正規母集団** $X_{1i} \sim N(\mu_1,\ \sigma_1^2)$**，** $i = 1, 2, \cdots, n_1$

第 2 グループ： 大きさ n_2 **の無作為標本，正規母集団** $X_{2j} \sim N(\mu_2,\ \sigma_2^2)$**，** $j = 1, 2, \cdots, n_2$