$f(x)$ and $F(x)$ are given by:

$$f(x) = (2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}x^2) = \frac{1}{\sigma}\phi(\frac{x}{\sigma}),$$
$$F(x) = \int_{-\infty}^{x}(2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}z^2)\mathrm{d}z = \Phi(\frac{x}{\sigma}).$$

[Review — Mean of Truncated Normal Random Variable:]

Let X be a normal random variable with mean μ and variance σ^2.

Consider $\mathrm{E}(X|X > a)$, where a is known.

The truncated distribution of X given $X > a$ is:

$$f(x|x > a) = \frac{(2\pi\sigma^2)^{-1/2}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)}{\displaystyle\int_{a}^{\infty}(2\pi\sigma^2)^{-1/2}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)\mathrm{d}x} = \frac{\frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})}{1-\Phi(\frac{a-\mu}{\sigma})}.$$

$$\mathrm{E}(X|X>a)=\int_a^\infty xf(x|x>a)\mathrm{d}x=\frac{\displaystyle\int_a^\infty x(2\pi\sigma^2)^{-1/2}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)\mathrm{d}x}{\displaystyle\int_a^\infty (2\pi\sigma^2)^{-1/2}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)\mathrm{d}x}$$

$$=\frac{\sigma\phi(\frac{a-\mu}{\sigma})+\mu\Big(1-\Phi(\frac{a-\mu}{\sigma})\Big)}{1-\Phi(\frac{a-\mu}{\sigma})}=\frac{\sigma\phi(\frac{a-\mu}{\sigma})}{1-\Phi(\frac{a-\mu}{\sigma})}+\mu,$$

which are shown below. The denominator is:

$$\int_a^\infty (2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\mathrm{d}x=\int_{\frac{a-\mu}{\sigma}}^\infty (2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z$$

$$=1-\int_{-\infty}^{\frac{a-\mu}{\sigma}} (2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z$$

$$=1-\Phi(\frac{a-\mu}{\sigma}),$$

where x is transformed into $z=\dfrac{x-\mu}{\sigma}$. $\qquad x>a \implies z=\dfrac{x-\mu}{\sigma}>\dfrac{a-\mu}{\sigma}$.

The numerator is:

$$
\begin{aligned}
&\int_a^\infty x(2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\mathrm{d}x \\
&= \int_{\frac{a-\mu}{\sigma}}^\infty (\sigma z+\mu)(2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z \\
&= \sigma\int_{\frac{a-\mu}{\sigma}}^\infty z(2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z + \mu\int_{\frac{a-\mu}{\sigma}}^\infty (2\pi)^{-1/2}\exp(-\frac{1}{2\sigma^2}z^2)\mathrm{d}z \\
&= \sigma\int_{\frac{1}{2}(\frac{a-\mu}{\sigma})^2}^\infty (2\pi)^{-1/2}\exp(-t)\mathrm{d}t + \mu\Big(1-\Phi(\frac{a-\mu}{\sigma})\Big) \\
&= \sigma\phi(\frac{a-\mu}{\sigma}) + \mu\Big(1-\Phi(\frac{a-\mu}{\sigma})\Big),
\end{aligned}
$$

where z is transformed into $t=\frac{1}{2}z^2$. $\quad z>\frac{a-\mu}{\sigma} \implies t=\frac{1}{2}z^2>\frac{1}{2}(\frac{a-\mu}{\sigma})^2$.

[End of Review]

Therefore, the conditional expectation of u_i given $X_i\beta + u_i > a$ is:

$$\mathrm{E}(u_i|X_i\beta + u_i > a) = \int_{a-X_i\beta}^{\infty} u_i \frac{f(u_i)}{1 - F(a - X_i\beta)} \mathrm{d}u_i = \int_{a-X_i\beta}^{\infty} \frac{u_i}{\sigma} \frac{\phi(\dfrac{u_i}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})} \mathrm{d}u_i$$

$$= \frac{\sigma\phi(\dfrac{a - X_i\beta}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})}.$$

Accordingly, the conditional expectation of y_i given $y_i > a$ is given by:

$$\mathrm{E}(y_i|y_i > a) = \mathrm{E}(y_i|X_i\beta + u_i > a) = \mathrm{E}(X_i\beta + u_i|X_i\beta + u_i > a)$$

$$= X_i\beta + \mathrm{E}(u_i|X_i\beta + u_i > a) = X_i\beta + \frac{\sigma\phi(\dfrac{a - X_i\beta}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})},$$

for $i = 1, 2, \cdots, n$.

Estimation:

MLE:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} \frac{f(y_i - X_i\beta)}{1 - F(a - X_i\beta)} = \prod_{i=1}^{n} \frac{1}{\sigma} \frac{\phi(\frac{y_i - X_i\beta}{\sigma})}{1 - \Phi(\frac{a - X_i\beta}{\sigma})}$$

is maximized with respect to β and σ^2.

Some Examples:

1. Buying a Car:

   $y_i = x_i\beta + u_i$, where y_i denotes expenditure for a car, and x_i includes income, price of the car, etc.

   Data on people who bought a car are observed.

   People who did not buy a car are ignored.

2. Working-hours of Wife:

   y_i represents working-hours of wife, and x_i includes the number of children, age, education, income of husband, etc.

3. Stochastic Frontier Model:

   $y_i = f(K_i, L_i) + u_i$, where y_i denotes production, K_i is stock, and L_i is amount of labor.

   We always have $y_i \le f(K_i, L_i)$, i.e., $u_i \le 0$.

   $f(K_i, L_i)$ is a maximum value when we input K_i and L_i.

Censored Regression Model or Tobit Model:

$$y_i = \begin{cases} X_i\beta + u_i, & \text{if } y_i > a, \\ a, & \text{otherwise.} \end{cases}$$

The probability which y_i takes a is given by:

$$P(y_i = a) = P(y_i \leq a) = F(a) \equiv \int_{-\infty}^{a} f(x)\mathrm{d}x,$$

where $f(\cdot)$ and $F(\cdot)$ denote the density function and cumulative distribution function of y_i, respectively.

Therefore, the likelihood function is:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} F(a)^{I(y_i=a)} \times f(y_i)^{1-I(y_i=a)},$$

where $I(y_i = a)$ denotes the indicator function which takes one when $y_i = a$ or zero otherwise.

When $u_i \sim N(0, \sigma^2)$, the likelihood function is:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} \Big(\int_{-\infty}^{a} (2\pi\sigma^2)^{-1/2} \exp(-\frac{1}{2\sigma^2}(y_i - X_i\beta)^2) \mathrm{d}y_i \Big)^{I(y_i=a)}$$
$$\times \Big((2\pi\sigma^2)^{-1/2} \exp(-\frac{1}{2\sigma^2}(y_i - X_i\beta)^2) \Big)^{1-I(y_i=a)},$$

which is maximized with respect to β and σ^2.