

STA 312s19 Formulas

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$$

$$E(X) \stackrel{\text{def}}{=} \sum_x x p_x(x) \text{ or } \int_{-\infty}^\infty x f_x(x) dx$$

$$Var(X) \stackrel{\text{def}}{=} E((X - \mu)^2)$$

$$L(\theta) = \prod_{i=1}^n p(y_i|\theta)$$

$$L(\theta) = \prod_{i=1}^n f(t_i|\theta)^{\delta_i} S(t_i|\theta)^{1-\delta_i}$$

$$\hat{\theta}_n \sim N(\theta, \frac{1}{nI(\theta)})$$

$$\hat{v}_n = 1 / -\ell''(\hat{\theta})$$

$$95\% \text{ CI: } \hat{\theta} \pm 1.96 \times S_{\hat{\theta}}$$

If $g: \mathbb{R} \rightarrow \mathbb{R}$

$$\hat{\theta}_n \sim N_k(\theta, \frac{1}{n} \mathcal{I}(\theta)^{-1})$$

$$\mathbf{H}(\theta) = \left[-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell(\theta) \right]$$

If $g: \mathbb{R}^k \rightarrow \mathbb{R}$

$$\dot{g}(\theta) = \left(\frac{\partial g}{\partial \theta_1}, \dots, \frac{\partial g}{\partial \theta_k} \right)$$

$$G^2 = -2 \log \left(\frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} \right) = -2 \log \left(\frac{L(\hat{\theta}_0)}{L(\hat{\theta})} \right)$$

$$S(t) \stackrel{\text{def}}{=} P(T > t) = 1 - F(t)$$

$$h(t) = \frac{f(t)}{S(t)}$$

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \text{ and } \Gamma(\frac{1}{2}) = \sqrt{\pi}$$

$$E(g(X)) = \sum_x g(x) p_x(x) \text{ or } \int_{-\infty}^\infty g(x) f_x(x) dx$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$L(\theta) = \prod_{i=1}^n f(y_i|\theta)$$

$$\ell(\theta) = \log L(\theta)$$

$$I(\theta) = -E \frac{\partial^2}{\partial \theta^2} \log f(X|\theta)$$

$$S_{\hat{\theta}} = \sqrt{\hat{v}_n}$$

$$Z_n = \frac{\hat{\theta} - \theta_0}{S_{\hat{\theta}}}$$

$$g(\hat{\theta}) \sim N(g(\theta), g'(\theta)^2 v_n)$$

$$\mathcal{I}(\theta) = \left[-E \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(Y; \theta) \right) \right]$$

$$\hat{\mathbf{V}}_n = \mathbf{H}(\hat{\theta})^{-1} \text{ estimates } \frac{1}{n} \mathcal{I}(\theta)^{-1}$$

$$g(\hat{\theta}) \sim N(g(\theta), \dot{g}(\theta) \mathbf{V}_n \dot{g}(\theta)^\top)$$

$$W_n = (\mathbf{L} \hat{\theta}_n - \mathbf{h})^\top (\mathbf{L} \hat{\mathbf{V}}_n \mathbf{L}^\top)^{-1} (\mathbf{L} \hat{\theta}_n - \mathbf{h})$$

$$h(t) \stackrel{\text{def}}{=} \lim_{\Delta \rightarrow 0} \frac{P(t < T < t + \Delta | T > t)}{\Delta}, \text{ where } \Delta > 0$$

$$S(t) = \exp\{-\int_0^t h(x) dx\} = e^{-H(t)}$$

| Distribution | Density | S(t) | E(T) | Median |
|-------------------------|--|--|--------------------------------------|--------------------------------------|
| Exponential | $f(t \lambda) = \lambda e^{-\lambda t}$ for $t \geq 0$ | $e^{-\lambda t}$ | $\frac{1}{\lambda}$ | $\frac{\log(2)}{\lambda}$ |
| Weibull | $f(t \alpha, \lambda) = \alpha \lambda (\lambda t)^{\alpha-1} \exp\{-(\lambda t)^\alpha\}$ for $t \geq 0$ | $e^{-(\lambda t)^\alpha}$ | $\frac{\Gamma(1+1/\alpha)}{\lambda}$ | $\frac{\log(2)^{1/\alpha}}{\lambda}$ |
| Gumbel $G(\mu, \sigma)$ | $f(y \mu, \sigma) = \frac{1}{\sigma} \exp\left\{\left(\frac{y-\mu}{\sigma}\right) - e^{\left(\frac{y-\mu}{\sigma}\right)}\right\}$ | $e^{-e^{\left(\frac{t-\mu}{\sigma}\right)}}$ | $\sigma \Gamma'(1) + \mu$ | $\sigma \log(\log(2)) + \mu$ |
| Log-normal | $f(t \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{(\log(t)-\mu)^2}{2\sigma^2}\right\} \frac{1}{t}$ for $t \geq 0$ | | $e^{\mu + \frac{1}{2}\sigma^2}$ | e^μ |

Log of standard exponential is Gumbel(0,1), also called the standard extreme value distribution.

If $Z \sim G(0, 1)$, MGF is $M_z(t) = \Gamma(t+1)$, $E(Z) = \Gamma'(1)$, $Var(Z) = \frac{\pi^2}{6}$, and $Y = \sigma Z + \mu \sim G(\mu, \sigma)$.

Log of Weibull with $\alpha = 1/\sigma$ and $\lambda = e^{-\mu}$ is Gumbel(μ, σ).

Kaplan-Meier estimate: Discrete time.

- p_j = the probability of surviving past time t_j , given survival to time t_{j-1} .
 $S(t_k) = \prod_{j=1}^k p_j$.
- d_j is the number of deaths at time t_j , and n_j is the number of individuals at risk before time t_j .
- $\hat{p}_j = \frac{n_j - d_j}{n_j}$, $\hat{S}(t_k) = \prod_{j=1}^k \hat{p}_j$, $\hat{S}(t) = \prod_{t_j \leq t} \hat{p}_j$.
- $\hat{S}(t) \sim N\left(S(t), S(t)^2 \sum_{t_j \leq t} \frac{1-p_j}{n_j p_j}\right)$.
- The standard error of $\hat{S}(t)$ is $\hat{S}(t) \sqrt{\sum_{t_j \leq t} \left(\frac{d_j}{n_j(n_j - d_j)}\right)}$.

Weibull Regression: $t_i = \exp\{\beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1}\} \cdot \epsilon_i^\sigma = e^{\mathbf{x}_i^\top \boldsymbol{\beta}} \epsilon_i^\sigma$, where $\epsilon_1 \sim \exp(1)$.

- $t_i \sim$ Weibull, with $\alpha = 1/\sigma$ and $\lambda = e^{-\mathbf{x}_i^\top \boldsymbol{\beta}}$.
- $E(t_i) = e^{\mathbf{x}_i^\top \boldsymbol{\beta}} \Gamma(\sigma + 1)$, $\text{Median}(t_i) = e^{\mathbf{x}_i^\top \boldsymbol{\beta}} \log(2)^\sigma$, $h_i(t) = \frac{1}{\sigma} \exp\{-\frac{1}{\sigma} \mathbf{x}_i^\top \boldsymbol{\beta}\} t^{\frac{1}{\sigma}-1}$.
- $S(t) = \exp\left\{-e^{-\frac{1}{\sigma} \mathbf{x}_i^\top \boldsymbol{\beta}} t^{\frac{1}{\sigma}}\right\}$

Log-normal Regression Regression: $t_i = \exp\{\beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1}\} \cdot \epsilon_i^\sigma = e^{\mathbf{x}_i^\top \boldsymbol{\beta}} \epsilon_i^\sigma$, where $\epsilon_1 \sim \text{Log-normal}(0,1)$.

- $t_i \sim \text{Log-normal}(\mu, \sigma^2)$, with $\mu = e^{\mathbf{x}_i^\top \boldsymbol{\beta}}$.
- $E(t_i) = e^{\mathbf{x}_i^\top \boldsymbol{\beta} + \frac{1}{2}\sigma^2}$, $\text{Median}(t_i) = e^{\mathbf{x}_i^\top \boldsymbol{\beta}}$.

Proportional Hazards Regression

- $h_i(t|\boldsymbol{\beta}) = h_0(t) e^{\mathbf{x}_i^\top \boldsymbol{\beta}}$.
- $\text{PL}(T) = \prod_{i=1}^D \left(\frac{e^{\mathbf{x}_{(i)}^\top \boldsymbol{\beta}}}{\sum_{j \in R_{(i)}} e^{\mathbf{x}_j^\top \boldsymbol{\beta}}} \right)$.
- $h_{i,j}(t_{i,j}) = h_0(t_{i,j}) \exp\{\sigma z_i + \mathbf{x}_{i,j}^\top \boldsymbol{\beta}\}$

This formula sheet was prepared by [Jerry Brunner](#), Department of Mathematical and Computational Sciences, University of Toronto Mississauga. It is licensed under a [Creative Commons Attribution - ShareAlike 3.0 Unported License](#). Use any part of it as you like and share the result freely. The \LaTeX source code is available from the course website:

<http://www.utstat.toronto.edu/~brunner/oldclass/312s19>