

## STA 442/2101 Formulas

$$M_y(t) = E(e^{yt})$$

$$M_{ay}(t) = M_y(at)$$

$$M_{y+a}(t) = e^{at} M_y(t)$$

$$M_{\sum_{i=1}^n y_i}(t) = \prod_{i=1}^n M_{y_i}(t)$$

$$y \sim N(\mu, \sigma^2) \text{ means } M_y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

$$y \sim \chi^2(\nu) \text{ means } M_y(t) = (1 - 2t)^{-\nu/2}$$

If  $W = W_1 + W_2$  with  $W_1$  and  $W_2$  independent,  $W \sim \chi^2(\nu_1 + \nu_2)$ ,  $W_2 \sim \chi^2(\nu_2)$  then  $W_1 \sim \chi^2(\nu_1)$

Columns of  $\mathbf{A}$  *linearly dependent* means there is a vector  $\mathbf{v} \neq \mathbf{0}$  with  $\mathbf{A}\mathbf{v} = \mathbf{0}$ .

Columns of  $\mathbf{A}$  *linearly independent* means that  $\mathbf{A}\mathbf{v} = \mathbf{0}$  implies  $\mathbf{v} = \mathbf{0}$ .

$\mathbf{A}$  *positive definite* means  $\mathbf{v}^\top \mathbf{A}\mathbf{v} > 0$  for all vectors  $\mathbf{v} \neq \mathbf{0}$ .

$$\Sigma = \mathbf{P}\Lambda\mathbf{P}^\top$$

$$\Sigma^{-1} = \mathbf{P}\Lambda^{-1}\mathbf{P}^\top$$

$$\Sigma^{1/2} = \mathbf{P}\Lambda^{1/2}\mathbf{P}^\top$$

$$\Sigma^{-1/2} = \mathbf{P}\Lambda^{-1/2}\mathbf{P}^\top$$

$$\text{cov}(\mathbf{w}) = E\{(\mathbf{w} - \boldsymbol{\mu}_w)(\mathbf{w} - \boldsymbol{\mu}_w)^\top\}$$

$$\text{cov}(\mathbf{w}, \mathbf{t}) = E\{(\mathbf{w} - \boldsymbol{\mu}_w)(\mathbf{t} - \boldsymbol{\mu}_t)^\top\}$$

$$\text{cov}(\mathbf{w}) = E\{\mathbf{w}\mathbf{w}^\top\} - \boldsymbol{\mu}_w\boldsymbol{\mu}_w^\top$$

$$\text{cov}(\mathbf{A}\mathbf{w}) = \mathbf{A}\text{cov}(\mathbf{w})\mathbf{A}^\top$$

If  $\mathbf{w} \sim N_p(\boldsymbol{\mu}, \Sigma)$ , then  $\mathbf{A}\mathbf{w} + \mathbf{c} \sim N_r(\mathbf{A}\boldsymbol{\mu} + \mathbf{c}, \mathbf{A}\Sigma\mathbf{A}^\top)$

and  $(\mathbf{w} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{w} - \boldsymbol{\mu}) \sim \chi^2(p)$

$L(\boldsymbol{\mu}, \Sigma) = |\Sigma|^{-n/2} (2\pi)^{-np/2} \exp -\frac{n}{2} \left\{ \text{tr}(\widehat{\Sigma}\Sigma^{-1}) + (\bar{\mathbf{y}} - \boldsymbol{\mu})^\top \Sigma^{-1}(\bar{\mathbf{y}} - \boldsymbol{\mu}) \right\}$ , where  $\widehat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})^\top$

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$y_i = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1} + \epsilon_i$$

$\epsilon_1, \dots, \epsilon_n$  independent  $N(0, \sigma^2)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} \sim N_p(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1})$$

$$\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{H}\mathbf{y}, \text{ where } \mathbf{H} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$$

$$\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}} = (\mathbf{I} - \mathbf{H})\mathbf{y}, \quad \mathbf{X}^\top \mathbf{e} = \mathbf{0}$$

$\widehat{\boldsymbol{\beta}}$  and  $\mathbf{e}$  are independent under normality.

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \widehat{y}_i)^2 + \sum_{i=1}^n (\widehat{y}_i - \bar{y})^2$$

$$SST = SSE + SSR \text{ and } R^2 = \frac{SSR}{SST}$$

$$\frac{SSE}{\sigma^2} = \frac{\mathbf{e}^\top \mathbf{e}}{\sigma^2} \sim \chi^2(n - p)$$

$$MSE = \frac{SSE}{n - p}$$

$$T = \frac{Z}{\sqrt{W/\nu}} \sim t(\nu)$$

$$F = \frac{W_1/\nu_1}{W_2/\nu_2} \sim F(\nu_1, \nu_2)$$

$$\text{Under } H_0 : \mathbf{L}\boldsymbol{\beta} = \mathbf{h}, F^* = \frac{(\mathbf{L}\widehat{\boldsymbol{\beta}} - \mathbf{h})^\top (\mathbf{L}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{L}^\top)^{-1} (\mathbf{L}\widehat{\boldsymbol{\beta}} - \mathbf{h})}{r MSE_F} = \frac{SSR_F - SSR_R}{r MSE_F} \sim F(r, n - p)$$

If  $\lim_{n \rightarrow \infty} E(T_n) = \theta$  and  $\lim_{n \rightarrow \infty} Var(T_n) = 0$ , then  $T_n \xrightarrow{p} \theta$

If  $\sqrt{n}(T_n - \mu) \xrightarrow{d} T \sim N(0, \sigma^2)$ , then  $\sqrt{n}(g(T_n) - g(\mu)) \xrightarrow{d} g'(\mu)T \sim N(0, g'(\mu)^2\sigma^2)$

If  $\mathbf{T}_n \xrightarrow{d} \mathbf{T}$  and  $\mathbf{Y}_n \xrightarrow{p} \mathbf{c}$ , then  $\begin{pmatrix} \mathbf{T}_n \\ \mathbf{Y}_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \mathbf{T} \\ \mathbf{c} \end{pmatrix}$   $\sqrt{n}(\bar{\mathbf{x}}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathbf{x} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$

Let  $g: \mathbb{R}^d \rightarrow \mathbb{R}^k$  etc. If  $\sqrt{n}(\mathbf{T}_n - \boldsymbol{\theta}) \xrightarrow{d} \mathbf{T}$ , then  $\sqrt{n}(g(\mathbf{T}_n) - g(\boldsymbol{\theta})) \xrightarrow{d} \dot{g}(\boldsymbol{\theta})\mathbf{T}$ , where  $\dot{g}(\boldsymbol{\theta}) = \left[ \frac{\partial g_i}{\partial \theta_j} \right]_{k \times d}$

$G^2 = -2 \log \left( \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} \right) = -2 \log \left( \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} \right)$   $W_n = (\mathbf{L}\hat{\boldsymbol{\theta}}_n - \mathbf{h})^\top (\mathbf{L}\hat{\mathbf{V}}_n\mathbf{L}^\top)^{-1} (\mathbf{L}\hat{\boldsymbol{\theta}}_n - \mathbf{h})$

$\log \left( \frac{\pi_i}{1 - \pi_i} \right) = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1}$   $\pi_i = \frac{e^{\beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1}}}{1 + e^{\beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1}}}$

$\log(\lambda_i) = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_{p-1} x_{i,p-1}$

$$\begin{aligned} \log \left( \frac{\pi_1}{\pi_3} \right) &= \beta_{0,1} + \beta_{1,1} x_1 + \dots + \beta_{p-1,1} x_{p-1} = L_1 & \pi_1 &= \frac{e^{L_1}}{1 + e^{L_1} + e^{L_2}} \\ \log \left( \frac{\pi_2}{\pi_3} \right) &= \beta_{0,2} + \beta_{1,2} x_1 + \dots + \beta_{p-1,2} x_{p-1} = L_2 & \pi_2 &= \frac{e^{L_2}}{1 + e^{L_1} + e^{L_2}} \\ & & \pi_3 &= \frac{1}{1 + e^{L_1} + e^{L_2}} \end{aligned}$$

```
> df = 1:12
> Critical_Value = qchisq(0.95,df)
> cbind(df,Critical_Value)
  df Critical_Value
[1,] 1          3.841459
[2,] 2          5.991465
[3,] 3          7.814728
[4,] 4          9.487729
[5,] 5         11.070498
[6,] 6         12.591587
[7,] 7         14.067140
[8,] 8         15.507313
[9,] 9         16.918978
[10,] 10        18.307038
[11,] 11        19.675138
[12,] 12        21.026070
```

---

This formula sheet was prepared by [Jerry Brunner](#), Department of Statistics, University of Toronto. 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 L<sup>A</sup>T<sub>E</sub>X source code is available from the course website: <http://www.utstat.toronto.edu/~brunner/oldclass/302f17>