

## 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  $\mathbf{v} \neq \mathbf{0}$ .

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

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

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

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

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

$$\mathbf{Y}_n = \sqrt{n}(\bar{\mathbf{Y}}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathbf{Y} \sim N(\mathbf{0}, \Sigma)$$

So asymptotically,  $\bar{\mathbf{Y}}_n \sim N(\boldsymbol{\mu}, \frac{1}{n}\Sigma)$

$$V(\mathbf{Y}) = E\{(\mathbf{Y} - \boldsymbol{\mu}_y)(\mathbf{Y} - \boldsymbol{\mu}_y)^\top\}$$

$$C(\mathbf{Y}, \mathbf{T}) = E\{(\mathbf{Y} - \boldsymbol{\mu}_y)(\mathbf{T} - \boldsymbol{\mu}_t)^\top\}$$

$$V(\mathbf{Y}) = E\{\mathbf{Y}\mathbf{Y}^\top\} - \boldsymbol{\mu}_y\boldsymbol{\mu}_y^\top$$

$$V(\mathbf{A}\mathbf{Y}) = \mathbf{A}V(\mathbf{Y})\mathbf{A}^\top$$

$$M_{\mathbf{Y}}(\mathbf{t}) = E(e^{\mathbf{t}^\top \mathbf{Y}})$$

$$M_{\mathbf{A}\mathbf{Y}}(\mathbf{t}) = M_{\mathbf{Y}}(\mathbf{A}^\top \mathbf{t})$$

$\mathbf{Y}_1$  and  $\mathbf{Y}_2$  are independent if and only if  $M_{(\mathbf{Y}_1, \mathbf{Y}_2)}(\mathbf{t}_1, \mathbf{t}_2) = M_{\mathbf{Y}_1}(\mathbf{t}_1)M_{\mathbf{Y}_2}(\mathbf{t}_2)$

$$M_{\mathbf{Y}+\mathbf{c}}(\mathbf{t}) = e^{\mathbf{t}^\top \mathbf{c}}M_{\mathbf{Y}}(\mathbf{t})$$

$$\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \Sigma) \text{ means } M_{\mathbf{Y}}(\mathbf{t}) = e^{\mathbf{t}^\top \boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^\top \Sigma \mathbf{t}}$$

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

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

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

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

$$W_n = (\mathbf{L}\widehat{\boldsymbol{\theta}}_n - \mathbf{h})^\top \left( \mathbf{L}\widehat{\mathbf{V}}_n\mathbf{L}^\top \right)^{-1} (\mathbf{L}\widehat{\boldsymbol{\theta}}_n - \mathbf{h})$$

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

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

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

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

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

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

$$\mathbf{e} = \mathbf{Y} - \widehat{\mathbf{Y}}$$

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

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

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

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

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

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

$$T = \frac{\mathbf{a}^\top \hat{\boldsymbol{\beta}} - \mathbf{a}^\top \boldsymbol{\beta}}{\sqrt{MSE \mathbf{a}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{a}}} \sim t(n-p)$$

$$F^* = \frac{SSR_F - SSR_R}{r MSE_F} = \left(\frac{n-p}{r}\right) \left(\frac{a}{1-a}\right) \sim F(r, n-p)$$

$$a = \frac{R_F^2 - R_R^2}{1 - R_R^2} = \frac{r F^*}{n-p+r F^*}$$

$$Z_i \stackrel{ind}{\sim} N(\mu_i, 1) \Rightarrow \sum_{i=1}^n Z_i^2 \sim \chi_{nc}^2(n, \lambda = \sum_{i=1}^n \mu_i^2)$$

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

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

$$T = \frac{Y_{n+1} - \mathbf{x}_{n+1}^\top \hat{\boldsymbol{\beta}}}{\sqrt{MSE (1 + \mathbf{x}_{n+1}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{x}_{n+1})}} \sim t(n-p)$$

$$F^* = \frac{(\mathbf{L} \hat{\boldsymbol{\beta}} - \mathbf{h})^\top (\mathbf{L} (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{C}^\top)^{-1} (\mathbf{L} \hat{\boldsymbol{\beta}} - \mathbf{h})}{r MSE} \sim F(r, n-p)$$

$$\lambda = \frac{(\mathbf{L} \boldsymbol{\beta} - \mathbf{h})^\top (\mathbf{L} (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{L}^\top)^{-1} (\mathbf{L} \boldsymbol{\beta} - \mathbf{h})}{\sigma^2}$$

$$\lambda = \frac{\sum_{k=1}^p n_k (\mu_k - \mu_\cdot)^2}{\sigma^2}, \text{ where } \mu_\cdot = \sum_{k=1}^p \frac{n_k}{n} \mu_k$$

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

```
> # Chi-squared critical values
> df = 1:6
> 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
```

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/302f14>