STA 2101/442 Formulas

$V(\mathbf{Y}) = E\{(\mathbf{Y}-\boldsymbol{\mu})(\mathbf{Y}-\boldsymbol{\mu})'\}$

$C(\mathbf{X},\mathbf{Y}) = E\left\{(\mathbf{X}-\boldsymbol{\mu}_x)(\mathbf{Y}-\boldsymbol{\mu}_y)'\right\}$

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

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

$\sqrt{n}(g(\overline{\mathbf{Y}}_n) - g(\boldsymbol{\mu})) \stackrel{d}{\rightarrow} \dot{\mathrm{g}}(\boldsymbol{\mu})\mathbf{Y}, \quad \dot{\mathrm{g}}(\mathbf{x}) = \left[\frac{\partial g_i}{\partial x_j}\right]_{k\times d}$

If $\mathbf{Y} \sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$, then $\mathbf{AY} \sim N_r(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$ and $(\mathbf{Y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu}) \sim \chi^2(p)$

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

$P(n_1,\ldots,n_c) = \binom{n}{n_1\,\cdots\,n_c}\pi_1^{n_1}\cdots\pi_c^{n_c}$

$L(\boldsymbol{\pi}) = \prod_{i=1}^n \pi_1^{y_{i,1}}\pi_2^{y_{i,2}}\cdots\pi_c^{y_{i,c}} = \pi_1^{n_1}\pi_2^{n_2}\cdots\pi_c^{n_c}$

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

$\boldsymbol{\mathcal{J}}_n(\boldsymbol{\theta}) = \left[\frac{1}{n}\sum_{i=1}^n -\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log f(Y_i;\boldsymbol{\theta})\right]$

$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \stackrel{d}{\rightarrow} \mathbf{T} \sim N_k\left(\mathbf{0}, \boldsymbol{\mathcal{I}}(\boldsymbol{\theta})^{-1}\right)$

$\widehat{\mathbf{V}}_n = \frac{1}{n}\boldsymbol{\mathcal{J}}_n(\widehat{\boldsymbol{\theta}}_n)^{-1} = \left(\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(\boldsymbol{\theta},\mathbf{Y})\right]_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n}\right)^{-1}$

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

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

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

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

$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

$\widehat{\mathbf{Y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{HY}$

$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

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

$\widehat{\boldsymbol{\beta}} \sim N_p\left(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right)$

$SSE/\sigma^2 \sim \chi^2(n-p)$

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

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

$F = \frac{(\mathbf{L}\widehat{\boldsymbol{\beta}}-\mathbf{h})'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\widehat{\boldsymbol{\beta}}-\mathbf{h})}{r\,MSE_F} = \frac{(SSR_F-SSR_R)/r}{MSE_F} = \left(\frac{n-p}{r}\right)\left(\frac{a}{1-a}\right)$, where $a = \frac{R_F^2-R_R^2}{1-R_R^2} = \frac{rF}{n-p+rF}$

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

```
> df = 1:8
> CriticalValue = qchisq(0.95,df)
> round(rbind(df,CriticalValue),3)
               [,1]  [,2]  [,3]  [,4]  [,5]   [,6]   [,7]   [,8]
df            1.000 2.000 3.000 4.000  5.00  6.000  7.000  8.000
CriticalValue 3.841 5.991 7.815 9.488 11.07 12.592 14.067 15.507
```

More large sample tools

1. Definitions (All quantities in boldface are vectors in $\mathbb{R}^m$ unless otherwise stated)
   - ⋆ $\mathbf{T}_n \stackrel{a.s.}{\rightarrow} \mathbf{T}$ means $P\{\omega : \lim_{n\rightarrow\infty} \mathbf{T}_n(\omega) = \mathbf{T}(\omega)\} = 1$.
   - ⋆ $\mathbf{T}_n \stackrel{P}{\rightarrow} \mathbf{T}$ means $\forall \epsilon > 0, \lim_{n\rightarrow\infty} P\{||\mathbf{T}_n - \mathbf{T}|| < \epsilon\} = 1$.
   - ⋆ $\mathbf{T}_n \stackrel{d}{\rightarrow} \mathbf{T}$ means for every continuity point $\mathbf{t}$ of $F_{\mathbf{T}}$, $\lim_{n\rightarrow\infty} F_{\mathbf{T}_n}(\mathbf{t}) = F_{\mathbf{T}}(\mathbf{t})$.
2. $\mathbf{T}_n \stackrel{a.s.}{\rightarrow} \mathbf{T} \Rightarrow \mathbf{T}_n \stackrel{P}{\rightarrow} \mathbf{T} \Rightarrow \mathbf{T}_n \stackrel{d}{\rightarrow} \mathbf{T}$.
3. If $\mathbf{a}$ is a vector of constants, $\mathbf{T}_n \stackrel{d}{\rightarrow} \mathbf{a} \Rightarrow \mathbf{T}_n \stackrel{P}{\rightarrow} \mathbf{a}$.
4. Strong Law of Large Numbers (SLLN): Let $\mathbf{X}_1, \ldots \mathbf{X}_n$ be independent and identically distributed random vectors with finite first moment, and let $\mathbf{X}$ be a general random vector from the same distribution. Then $\overline{\mathbf{X}}_n \stackrel{a.s.}{\rightarrow} E(\mathbf{X})$.
5. Central Limit Theorem: Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. random vectors with expected value vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then $\sqrt{n}(\overline{\mathbf{X}}_n - \boldsymbol{\mu})$ converges in distribution to a multivariate normal with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$.
6. Slutsky Theorems for Convergence in Distribution:
   (a) If $\mathbf{T}_n \in \mathbb{R}^m$, $\mathbf{T}_n \stackrel{d}{\rightarrow} \mathbf{T}$ and if $f : \mathbb{R}^m \rightarrow \mathbb{R}^q$ (where $q \leq m$) is continuous except possibly on a set C with $P(\mathbf{T} \in C) = 0$, then $f(\mathbf{T}_n) \stackrel{d}{\rightarrow} f(\mathbf{T})$.
   (b) If $\mathbf{T}_n \stackrel{d}{\rightarrow} \mathbf{T}$ and $(\mathbf{T}_n - \mathbf{Y}_n) \stackrel{P}{\rightarrow} 0$, then $\mathbf{Y}_n \stackrel{d}{\rightarrow} \mathbf{T}$.
   (c) If $\mathbf{T}_n \in \mathbb{R}^d$, $\mathbf{Y}_n \in \mathbb{R}^k$, $\mathbf{T}_n \stackrel{d}{\rightarrow} \mathbf{T}$ and $\mathbf{Y}_n \stackrel{P}{\rightarrow} \mathbf{c}$, then
   $$\left(\begin{array}{c} \mathbf{T}_n \\ \mathbf{Y}_n \end{array} \right) \stackrel{d}{\rightarrow} \left(\begin{array}{c} \mathbf{T} \\ \mathbf{c} \end{array} \right)$$
7. Slutsky Theorems for Convergence in Probability:
   (a) If $\mathbf{T}_n \in \mathbb{R}^m$, $\mathbf{T}_n \stackrel{P}{\rightarrow} \mathbf{T}$ and if $f : \mathbb{R}^m \rightarrow \mathbb{R}^q$ (where $q \leq m$) is continuous except possibly on a set C with $P(\mathbf{T} \in C) = 0$, then $f(\mathbf{T}_n) \stackrel{P}{\rightarrow} f(\mathbf{T})$.
   (b) If $\mathbf{T}_n \stackrel{P}{\rightarrow} \mathbf{T}$ and $(\mathbf{T}_n - \mathbf{Y}_n) \stackrel{P}{\rightarrow} 0$, then $\mathbf{Y}_n \stackrel{P}{\rightarrow} \mathbf{T}$.
   (c) If $\mathbf{T}_n \in \mathbb{R}^d$, $\mathbf{Y}_n \in \mathbb{R}^k$, $\mathbf{T}_n \stackrel{P}{\rightarrow} \mathbf{T}$ and $\mathbf{Y}_n \stackrel{P}{\rightarrow} \mathbf{Y}$, then
   $$\left(\begin{array}{c} \mathbf{T}_n \\ \mathbf{Y}_n \end{array} \right) \stackrel{P}{\rightarrow} \left(\begin{array}{c} \mathbf{T} \\ \mathbf{Y} \end{array} \right)$$
8. Delta Method (Theorem of Cramér, Ferguson p. 45): Let $g : \mathbb{R}^d \rightarrow \mathbb{R}^k$ be such that the elements of $\dot{\mbox{g}}(\mathbf{x}) = \left[\frac{\partial g_i}{\partial x_j} \right]_{k \times d}$ are continuous in a neighborhood of $\boldsymbol{\theta} \in \mathbb{R}^d$. If $\mathbf{T}_n$ is a sequence of d-dimensional random vectors such that $\sqrt{n}(\mathbf{T}_n - \boldsymbol{\theta}) \stackrel{d}{\rightarrow} \mathbf{T}$, then $\sqrt{n}(g(\mathbf{T}_n) - g(\boldsymbol{\theta})) \stackrel{d}{\rightarrow} \dot{\mbox{g}}(\boldsymbol{\theta})\mathbf{T}$. In particular, if $\sqrt{n}(\mathbf{T}_n - \boldsymbol{\theta}) \stackrel{d}{\rightarrow} \mathbf{T} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, then $\sqrt{n}(g(\mathbf{T}_n) - g(\boldsymbol{\theta})) \stackrel{d}{\rightarrow} \mathbf{Y} \sim N(\mathbf{0}, \dot{\mbox{g}}(\boldsymbol{\theta})\boldsymbol{\Sigma}\dot{\mbox{g}}(\boldsymbol{\theta})')$.
