STA 302 Formulas[1] .

$E(X) \stackrel{def}{=} \sum_x x\, p_{_X}(x)$

$E(X) \stackrel{def}{=} \int_{-\infty}^{\infty} x f_{_X}(x)\, dx$

$E(g(X)) = \sum_x g(x)\, p_{_X}(x)$

$E(g(\mathbf{X})) = \sum_{x_1} \cdots \sum_{x_p} g(x_1,\ldots,x_p)\, p_{\mathbf{x}}(x_1,\ldots,x_p)$

$E(g(X)) = \int_{-\infty}^{\infty} g(x)\, f_{_X}(x)\, dx$

$E(g(\mathbf{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1,\ldots,x_p)\, f_{\mathbf{x}}(x_1,\ldots,x_p)\, dx_1 \ldots dx_p$

$E(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i E(X_i)$

$Var(X) \stackrel{def}{=} E\left((X-\mu_{_X})^2 \right) = E(X^2) - [E(X)]^2$

$Cov(X,Y) \stackrel{def}{=} E\left((X-\mu_{_X})(Y-\mu_{_Y}) \right)$

$Cov(X,Y) = E(XY) - E(X)E(Y)$

$Corr(X,Y) \stackrel{def}{=} \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

$Cov\left(\sum_{i=1}^n a_i X_i \,,\, \sum_{j=1}^m b_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(X_i, Y_j)$

$M_{_X}(t) = E(e^{Xt})$

$M_{_{aX}}(t) = M_{_X}(at)$

$M_{_{X+a}}(t) = e^{at} M_{_X}(t)$

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

$X \sim N(\mu,\sigma^2)$ means $M_{_X}(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$X \sim \chi^2(\nu)$ means $M_{_X}(t) = (1-2t)^{-\nu/2}$

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 $W_1 \sim \chi^2(\nu_1)$

If $x_1,\ldots,x_n$ is a random sample from a Normal(μ,σ^2) distribution, then $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, and

- $\widehat{\mu} = \overline{x}_n$ and $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x}_n)^2 = \left(\frac{n-1}{n}\right) s^2$, where $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x}_n)^2$
- $\overline{x}_n$ and s^2 are independent.
- $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$.
- $t = \frac{\sqrt{n}(\bar{s}-\mu)}{s} \sim t(n-1)$

$\left(\begin{array}{cc} a & b \\ c & d \end{array} \right)^{-1} = \frac{1}{ad-bc} \left(\begin{array}{rr} d & -b \\ -c & a \end{array} \right)$

$\mathbf{AB} = \left[\sum_k a_{i,k} b_{k,j} \right]$

$tr(\mathbf{AB}) = tr(\mathbf{BA})$

The square matrix $\mathbf{A}$ has an eigenvalue equal to λ with corresponding eigenvector $\mathbf{x} \neq \mathbf{0}$ if $\mathbf{Ax} = \lambda\mathbf{x}$.

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

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

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

$\boldsymbol{\Sigma} = \mathbf{CDC}'$

$\boldsymbol{\Sigma}^{-1} = \mathbf{CD}^{-1}\mathbf{C}'$

$\boldsymbol{\Sigma}^{1/2} = \mathbf{CD}^{1/2}\mathbf{C}'$

$\boldsymbol{\Sigma}^{-1/2} = \mathbf{CD}^{-1/2}\mathbf{C}'$

$cov(\mathbf{y}) = E\left\{ (\mathbf{y}-\boldsymbol{\mu}_y)(\mathbf{y}-\boldsymbol{\mu}_y)' \right\}$

$cov(\mathbf{y},\mathbf{t}) = E\left\{ (\mathbf{y}-\boldsymbol{\mu}_y)(\mathbf{t}-\boldsymbol{\mu}_t)' \right\}$

$cov(\mathbf{y}) = E\{\mathbf{yy}'\} - \boldsymbol{\mu}_y\boldsymbol{\mu}_y'$

$cov(\mathbf{Ay}) = \mathbf{A}cov(\mathbf{y})\mathbf{A}'$

$cov(\mathbf{Ay},\mathbf{By}) = \mathbf{A}cov(\mathbf{y})\mathbf{B}'$

[1] This formula sheet was prepared by Jerry Brunner, Department of Statistical Sciences, University of Toronto. The LaTeX source code is available from the course website: http://www.utstat.toronto.edu/~brunner/oldclass/302f20

$M_{\mathbf{v}}(\mathbf{t}) = E(e^{\mathbf{t}'\mathbf{v}})$

$M_{\mathbf{Av}}(\mathbf{t}) = M_{\mathbf{v}}(\mathbf{A}'\mathbf{t})$

$M_{\mathbf{v}+\mathbf{c}}(\mathbf{t}) = e^{\mathbf{t}'\mathbf{c}} M_{\mathbf{v}}(\mathbf{t})$

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

$\mathbf{v} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means $M_{\mathbf{v}}(\mathbf{t}) = e^{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$

For the multivariate normal, zero covariance implies independence.

If $\mathbf{v} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{Av} + \mathbf{c} \sim N_q(\mathbf{A}\boldsymbol{\mu} + \mathbf{c}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$,

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

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

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

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

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

$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i$

with $E(\epsilon_i) = 0$, $Var(\epsilon_i) = \sigma^2$, $Cov(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$

$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $E(\boldsymbol{\epsilon}) = \mathbf{0}$, $cov(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}_n$

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

$\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{Hy}$, where $\mathbf{H} = (h_{ij}) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

$\widehat{\boldsymbol{\epsilon}} = \mathbf{y} - \widehat{\mathbf{y}} = (\mathbf{I} - \mathbf{H})\mathbf{y}$ $\qquad \mathbf{X}'\widehat{\boldsymbol{\epsilon}} = \mathbf{0}$

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

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

$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$

$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i$
$\epsilon_1, \ldots, \epsilon_n$ independent $N(0, \sigma^2)$

$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \sim N_{k+1}\left(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right)$

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

$s^2 = \frac{\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}}}{n-k-1} = \frac{SSE}{n-k-1} = MSE$

$\frac{SSE}{\sigma^2} = \frac{\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}}}{\sigma^2} \sim \chi^2(n-k-1)$

$t = \frac{z}{\sqrt{w/\nu}} \sim t(\nu)$

$F = \frac{w_1/\nu_1}{w_2/\nu_2} \sim F(\nu_1, \nu_2)$

$t = \frac{\mathbf{a}'\widehat{\boldsymbol{\beta}} - \mathbf{a}'\boldsymbol{\beta}}{\sqrt{MSE\,\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}}} \sim t(n-k-1)$

$\mathbf{a}'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE\,\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}}$

$F^* = \frac{(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})}{q\,MSE} \overset{H_0}{\sim} F(q, n-k-1)$

$F^* = \frac{SSR(full) - SSR(reduced)}{q\,MSE(full)} = \left(\frac{n-k-1}{q}\right)\left(\frac{p}{1-p}\right)$

$p = \frac{R^2(full) - R^2(reduced)}{1 - R^2(reduced)}$

$p = \frac{qF^*}{qF^* + n - k - 1}$

$t = \frac{y_0 - \mathbf{x}_0'\widehat{\boldsymbol{\beta}}}{\sqrt{MSE\,(1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)}} \sim t(n-k-1)$

$\mathbf{x}_0'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE\,(1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)}$

$t_i = \frac{y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_{(i)}}{\sqrt{MSE_{(i)}(1 + \mathbf{x}_i'(\mathbf{X}_{(i)}'\mathbf{X}_{(i)})^{-1}\mathbf{x}_i)}} \sim t(n-k-2)$

$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{V})$

$\widehat{\boldsymbol{\beta}}_{gls} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$

```
qf(0.95,df1=6,df2=122) # Critical value for F, not in any table
qt(0.975,df=122)       $ Critical value for t

XpX= t(X)%*%X
XpXinv = solve(XpX)

install.packages("readxl") # Only need to do this once. There are no dependencies.
library(readxl) # Load the package
hungry = read_excel("Diet.xlsx")

math = read.table("http://www.utstat.toronto.edu/~brunner/data/legal/mathtest.txt")
head(math)
colnames(math) = c("ID","HScalcMark","PreCalcScore","CalcScore","UnivCalcMark")
summary(math)
cor(math)
attach(math)
mathmod = lm(UnivCalcMark ~ HScalcMark+PreCalcScore+CalcScore, data = math)
cellmeans = lm(lper100k ~ 0+Cntry+weight+length)
summary(mathmod)
betahat = coefficients(mathmod)
epsilonhat = residuals(mathmod)
plot(HScalcMark,epsilonhat)
yhat = fitted.values(mathmod)
hii = hatvalues(mathmod)
cd = cooks.distance(mathmod); summary(cd)
MSE.XpXinv = vcov(mathmod)

a = as.matrix(c(0,0,-1,1))
se = sqrt( t(a)%*%MSE.XpXinv%*%a ) # Standard error of the difference
me95 = as.numeric( t.025*se ) # Now me95 is different
estdiff = as.numeric( t(a) %*% betahat ); estdiff
Lower95 = estdiff - me95; Upper95 = estdiff + me95; c(Lower95, Upper95)

ti = rstudent(mathmod) # Studentized deleted residuals
alpha = 0.05; a = alpha/200; bcrit = qt(1-a/2,dfe-1); bcrit

# H0: C beta = t        ftest(model, C, t=0)
source("http://www.utstat.utoronto.ca/~brunner/Rfunctions/ftest.txt")
C1 = rbind( c(0,1,0,0),
            c(0,0,1,0),
            c(0,0,0,1) )
ftest(mod,C1)
anova(reducedmodel, fullmodel)

c1 = numeric(n); c1[Cntry=='Europ'] = 1; table(c1,Cntry)
c2 = numeric(n); c2[Cntry=='Japan'] = 1; table(c2,Cntry)
c3 = numeric(n); c3[Cntry=='US'] = 1;    table(c3,Cntry)

wc1 = weight*c1; wc2 = weight*c2
uneqslope = lm(lper100k ~ weight+c1+c2+wc1+wc2)

Cntry = factor(Cntry)
contrasts(Cntry)
Country  = Cntry
contrasts(Country) = contr.treatment(3,base=3)
```