Some formulæ for ENGI 4421 Probability and Statistics

On these pages are many more formulae than will fit comfortably on your allocation of two sheets in the final examination. Examine these suggestions together with your own notes, in order to construct a more concise set of formula sheets that will work best for *you*.

Descriptive Statistics

median = middle value in ordered set

mode = most frequently occurring value

$$\text{mean} = \bar{x} = \frac{1}{n}\sum x$$

$$s^2 = \frac{1}{n-1}\sum\left(x-\bar{x}\right)^2 \quad \text{or} \quad s^2 = \frac{n\sum x^2 - \left(\sum x\right)^2}{n\left(n-1\right)}$$

From a frequency table (f_i occurrences of distinct value x_i observed):

$$\text{mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}, \qquad \left(n = \sum f_i\right)$$

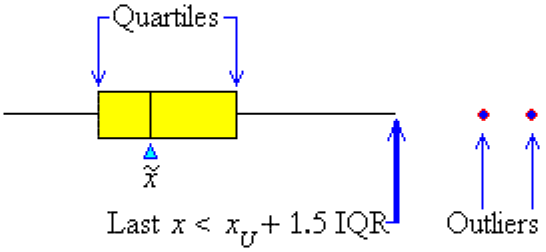
$$s^2 = \frac{1}{n-1}\sum f_i\left(x_i-\bar{x}\right)^2 \quad \text{or} \quad s^2 = \frac{n\sum f_i x_i^2 - \left(\sum f_i x_i\right)^2}{n\left(n-1\right)}$$

Quartiles: in ordered set $\{ x_1, x_2, \ldots, x_n \}$ are $x_{(n+1)/4}$ and $x_{3(n+1)/4}$

(interpolate as necessary)

Histogram: (area of a bar) = (rel. freq. of that interval)

Boxplot:



Decision Tree

[See examples in problem sets, past exams and additional exercises.]

Counting Techniques:

Number of distinct ways of selecting r objects from n objects:

with replacement, order of selection matters = n^r

without replacement, order of selection matters = ${}^nP_r = n! / (n-r)!$

without replacement, order of selection doesn't matter = ${}^nC_r = n! / (r!\,(n-r)!)$

${}^nP_0 = {}^nC_0 = {}^nC_n = 1, \qquad {}^nP_1 = {}^nC_1 = {}^nC_{n-1} = n, \qquad {}^nP_{n-1} = {}^nP_n = n!, \qquad {}^nC_{n-r} = {}^nC_r$

Number of distinct ways to partition n objects into piles of $r_1, r_2, \ldots, r_k$ objects:

$$\begin{pmatrix} n \\ r_1\; r_2 \cdots r_k \end{pmatrix} = \frac{n!}{r_1!\, r_2! \cdots r_k!}$$

Laws of Probability

Odds [on] = r and probability = p:

$$r = \frac{p}{1-p}, \quad p = \frac{r}{1+r}$$

deMorgan's laws: $\sim(A \cup B) = \sim A \cap \sim B$ and $\sim(A \cap B) = \sim A \cup \sim B$

General addition law of probability: $\mathrm{P}[A \vee B] = \mathrm{P}[A] + \mathrm{P}[B] - \mathrm{P}[A \wedge B]$

General multiplication law of probability: $\mathrm{P}[A \wedge B] = \mathrm{P}[A] \times \mathrm{P}[B \mid A] = \mathrm{P}[B] \times \mathrm{P}[A|B]$

Events A, B are independent iff $\mathrm{P}[A \wedge B] = \mathrm{P}[A] \times \mathrm{P}[B]$

Events A, B are incompatible (mutually exclusive) iff $\mathrm{P}[A \wedge B] = 0$

If $\{ E_1, E_2, \ldots, E_n \}$ is a partition, then **Total Probability Law:**

$$\mathrm{P}[A] = \sum_i \mathrm{P}[A \cap E_i] = \sum_i \mathrm{P}[A \mid E_i]\,\mathrm{P}[E_i]$$

and **Bayes' Theorem:**

$$\mathrm{P}[E_k \mid A] = \frac{\mathrm{P}[A \mid E_k]\,\mathrm{P}[E_k]}{\sum_i \mathrm{P}[A \mid E_i]\,\mathrm{P}[E_i]}$$

Conditions for $p(x)$ to be a probability mass function (p.m.f.):

$$p(x) \geq 0 \quad \forall x \qquad \text{and} \qquad \sum_{\text{all } x} p(x) = 1 \; (\text{coherence condition})$$

Conditions for $f(x)$ to be a probability density function (p.d.f.):

$$f(x) \geq 0 \quad \forall x \qquad \text{and} \qquad \int_{-\infty}^{\infty} f(x)\,dx = 1 \; (\text{coherence condition})$$

Cumulative Distribution Function (c.d.f.):

Discrete: Continuous:

$$F(x) = \sum_{y \leq x} p(x) \qquad F(x) = \int_{-\infty}^{x} f(x)\,dx \qquad \mathrm{P}[a < X \leq b] = F(b) - F(a)$$

(staircase) (ogive)

Median: $F(\tilde{\mu}) = \frac{1}{2}$ Lower quartile: $F(x_L) = \frac{1}{4}$ Upper quartile: $F(x_U) = \frac{3}{4}$

Expected value:

$$\mu = \mathrm{E}\left[X\right] = \sum_x x \cdot p(x) \quad \text{or} \quad \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$\mathrm{E}\left[h(X)\right] = \sum_x h(x)\cdot p(x) \quad \text{or} \quad \int_{-\infty}^{\infty} h(x)\, f(x)\, dx$$

Variance: $\sigma^2 = \mathrm{V}\left[X\right] = \mathrm{E}\left[(X-\mu)^2\right] = \mathrm{E}\left[X^2\right] - \left(\mathrm{E}\left[X\right]\right)^2$

Joint Probability Mass Functions and Joint Probability Density Functions:

$p(x,y) = \mathrm{P}\left[(X=x)\wedge(Y=y)\right]$

Marginal p.m.f.'s:

$p_X(x) = \sum_y p(x,y)$ and $p_Y(y) = \sum_x p(x,y)$

Marginal p.d.f.'s:

$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$

and $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$

Conditional p.m.f.'s:

$$p_{Y|X}(y\,|\,x) = \frac{p(x,y)}{p_X(x)}, \quad p_{X|Y}(x\,|\,y) = \frac{p(x,y)}{p_Y(y)}$$

Conditional p.d.f.'s: $f_{Y|X}(y\,|\,x) = \dfrac{f(x,y)}{f_X(x)}$ and $f_{X|Y}(x\,|\,y) = \dfrac{f(x,y)}{f_Y(y)}$

Covariance: $\mathrm{Cov}\left[X,Y\right] = \mathrm{E}\left[(X-\mu_X)(Y-\mu_Y)\right] = \mathrm{E}\left[XY\right] - \mathrm{E}\left[X\right]\mathrm{E}\left[Y\right]$

$\mathrm{V}[X] = \mathrm{Cov}[X,X] \geq 0$

Correlation: $\rho_{X,Y} = \dfrac{\mathrm{Cov}\left[X,Y\right]}{\sqrt{\mathrm{V}\left[X\right]\cdot\mathrm{V}\left[Y\right]}}; \qquad -1 \leq \rho \leq +1$

Independence $\Rightarrow \rho = 0$; (but $\rho = 0 \not\Rightarrow$ independence)

Combinations of Random Quantities:

$$\mathrm{E}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i\, \mathrm{E}\left[X_i\right], \qquad \mathrm{V}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\, a_j\, \mathrm{Cov}\left[X_i, X_j\right]$$

$\mathrm{E}[aX+bY] = a\,\mathrm{E}[X] + b\,\mathrm{E}[Y]$ and
$\mathrm{V}[aX+bY] = a^2\,\mathrm{V}[X] + 2ab\,\mathrm{Cov}[X,Y] + b^2\,\mathrm{V}[Y]$
for all constants a, b and all random quantities X, Y.
Special case, when X and Y are independent (or at least uncorrelated):
$\mathrm{E}[X \pm Y] = \mathrm{E}[X] \pm \mathrm{E}[Y]$ and $\mathrm{V}[X \pm Y] = \mathrm{V}[X] + \mathrm{V}[Y]$

$$\mathrm{E}\left[\overline{X}\right] = \mu, \qquad \mathrm{V}\left[\overline{X}\right] = \frac{\sigma^2}{n}$$

Propagation of Error

Sample size n leads to estimate $\left(\bar{x} \pm \frac{s}{\sqrt{n}} \right)$

Pooling $\left(x_1 \pm \sigma_1\right)$ with $\left(x_2 \pm \sigma_2\right)$: The most precise estimate is $\left(x \pm \sigma\right)$ where

$$x = c_1 x_1 + c_2 x_2 \,, \quad \sigma^2 = {c_1}^2 {\sigma_1}^2 + {c_2}^2 {\sigma_2}^2 \,, \quad c_1 = \frac{{\sigma_2}^2}{{\sigma_1}^2 + {\sigma_2}^2} \quad \text{and} \quad c_2 = \frac{{\sigma_1}^2}{{\sigma_1}^2 + {\sigma_2}^2}$$

If not independent, $Y = \sum_{i=1}^{n} a_i X_i \quad \Rightarrow \quad \sigma_Y \le |a_1|\sigma_1 + |a_2|\sigma_2 + \ldots + |a_n|\sigma_n$

Non-linear: $U = U\left(X\right) \quad \Rightarrow \quad \sigma_U \approx \left| \left(\left. \frac{dU}{dX} \right|_{X = x_{\text{obs}}} \right) \right| \sigma_X$

Multivariate and independent:

$$U = U\left(X_1, X_2, \ldots, X_n\right) \quad \Rightarrow \quad \sigma_U \approx \sqrt{\left(\frac{\partial U}{\partial X_1}\right)^2 {\sigma_1}^2 + \left(\frac{\partial U}{\partial X_2}\right)^2 {\sigma_2}^2 + \ldots + \left(\frac{\partial U}{\partial X_n}\right)^2 {\sigma_n}^2}$$

Multivariate and not independent:

$$\sigma_U \le \left|\frac{\partial U}{\partial X_1}\right| \sigma_1 + \left|\frac{\partial U}{\partial X_2}\right| \sigma_2 + \ldots + \left|\frac{\partial U}{\partial X_n}\right| \sigma_n$$

Bernoulli distribution:

$$p(x) = \begin{cases} p & (x=1) \\ 1-p & (x=0) \end{cases}, \qquad \mu = p \,, \qquad \sigma^2 = p(1-p)$$

Discrete Uniform distribution:

$$p(x) = \begin{cases} \frac{1}{n} & \left(x = x_1, x_2, \cdots, x_n\right) \\ 0 & (\text{otherwise}) \end{cases}, \qquad \mu = \frac{1}{n}\sum x$$

Binomial distribution bin(*n*, *p*):

$$\mathrm{P}\left[X = x\right] = \mathrm{b}\left(x; n, p\right) = {}^{n}C_{x}\, p^{x} \left(1-p\right)^{n-x}, \quad \mu = np, \quad \sigma^2 = np\left(1-p\right)$$

$$\mathrm{P}\left[X \le x\right] = \mathrm{B}\left(x; n, p\right) = \mathrm{b}\left(0; n, p\right) + \mathrm{b}\left(1; n, p\right) + \ldots + \mathrm{b}\left(x; n, p\right)$$

Four conditions for X to be binomial:

1) Each trial has a complementary pair of outcomes;
2) p = P[success] is the same for all trials;
3) Trials are all independent of each other;
4) The number of trials, n, is fixed.

Hypergeometric distribution *H*(*n*, *R*, *N*):

$$\mathrm{P}\left[X=x\right] = p(x) = \frac{{}^{R}C_{x} \times {}^{N-R}C_{n-x}}{{}^{N}C_{n}}$$ (may use binomial instead if $n < 5\%$ of *N*)

Poisson distribution Poisson(λ):

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad (x=0,1,2,\cdots); \quad \sigma^2 = \mu = \lambda$$

Continuous Uniform distribution U(*a*, *b*):

$$f(x) = \frac{1}{b-a} \quad (a \le x \le b); \quad \mathrm{P}\left[X \le x\right] = F(x) = \begin{cases} 0 & (x<a) \\ \dfrac{x-a}{b-a} & (a \le x \le b) \\ 1 & (x>b) \end{cases}$$

$$\mu = \frac{a+b}{2}, \qquad \sigma^2 = \frac{(b-a)^2}{12}$$

Normal distribution N(μ, σ^2):

To convert $X \sim \mathrm{N}\left(\mu,\sigma^2\right)$ to $Z \sim \mathrm{N}(0,1)$:

$$Z = \frac{X-\mu}{\sigma}; \qquad \text{then} \quad \mathrm{P}\left[a<X<b\right] = \mathrm{P}\left[z_a < Z < z_b\right] = \Phi\left(z_b\right) - \Phi\left(z_a\right)$$

Exponential distribution:

$$f(x) = \begin{cases} 0 & (x<0) \\ \lambda\, e^{-\lambda x} & (x \ge 0) \end{cases}; \quad \mathrm{P}\left[X \le x\right] = F(x) = \begin{cases} 0 & (x<0) \\ 1-e^{-\lambda x} & (x \ge 0) \end{cases}$$

$$\mathrm{P}\left[X \ge x\right] = e^{-\lambda x} \quad (x \ge 0); \quad \sigma = \mu = \frac{1}{\lambda}$$

Gamma distribution (Erlang if $r \in \mathbb{N}$):

$$f(t) = \frac{\lambda^r t^{r-1}}{\Gamma(r)} e^{-\lambda t} \quad (t>0), \quad \mu = \frac{r}{\lambda}, \quad \sigma^2 = \frac{r}{\lambda^2}; \text{ for Erlang, } \Gamma(r) = (r-1)!$$

Central Limit Theorem:

For any random quantity X with $\mathrm{E}[X] = \mu,\ \mathrm{V}[X] = \sigma^2$ and sufficiently large sample size n

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{and} \quad Z = \frac{\bar{X} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} \sim \mathrm{N}(0,1)$$

Bayesian Confidence Intervals:

If prior information is $X \sim \mathrm{N}\left(\mu_o, \sigma_o^{\,2}\right)$ and

data from random sample: size n, mean = $\bar{x}$, standard deviation = s, then
Calculate weights

$$w_o = \frac{1}{\sigma_o^{\,2}}, \quad w_d = \frac{n}{s^2} \quad \rightarrow \quad w^* = w_o + w_d$$

and updated mean and variance

$$\mu^* = \frac{w_d \bar{x} + w_o \mu_o}{w^*}, \qquad \left(\sigma^2\right)^* = \frac{1}{w^*} \quad \rightarrow \quad \text{Posterior dist'n } \mathrm{N}\left(\mu^*, \left(\sigma^2\right)^*\right)$$

Then the $(1-\alpha)\times 100\%$ confidence interval estimate for μ is

$$\mu^* - t_{\alpha/2,\, n-1} \cdot \sigma^* < \mu < \mu^* + t_{\alpha/2,\, n-1} \cdot \sigma^*$$

[***Note*****:** if σ is known, then it replaces s , and t becomes z.
Otherwise, the true number of degrees of freedom on t is actually a number between $n - 1$ and ∞. The interval shown above is a conservative approximation.]

Classical Confidence Intervals on μ:

$$\sigma_o^{\,2} \to \infty \quad \Rightarrow \quad w_o \to 0 \quad \Rightarrow \quad \mu^* = \bar{x} \quad \text{and} \quad \left(\sigma^2\right)^* = \frac{s^2}{n}$$

Then the $(1-\alpha)\times 100\%$ confidence interval estimate for μ is

$$\bar{x} - t_{\alpha/2,\, n-1} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2,\, n-1} \cdot \frac{s}{\sqrt{n}}$$

Classical Confidence Intervals on *p*:

$$p^* \pm z_{\alpha/2} \sqrt{\frac{p^* q^*}{n^*}} \quad \text{where} \quad p^* = \frac{x^*}{n^*} = \frac{x+2}{n+4}$$

Classical Confidence Intervals on $\mu_1 - \mu_2$ (large sample sizes & independent):

$$\left(\bar{x}_1 - \bar{x}_2\right) - z_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu < \left(\bar{x}_1 - \bar{x}_2\right) + z_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Classical Confidence Intervals on $p_1 - p_2$:

$$p_X^* - p_Y^* \pm z_{\alpha/2} \sqrt{\frac{p_X^* q_X^*}{n_X^*} + \frac{p_Y^* q_Y^*}{n_Y^*}}$$

where $x^* = x + 1,\ y^* = y + 1,\ n_X^* = n_X + 2$ and $n_Y^* = n_Y + 2$

Classical Confidence Intervals on $\mu_1 - \mu_2$ (small sample sizes & independent):

$$\left(\bar{x}_1 - \bar{x}_2\right) - t_{\alpha/2,\nu} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \;<\; \mu \;<\; \left(\bar{x}_1 - \bar{x}_2\right) + t_{\alpha/2,\nu} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{where} \quad \nu = \text{INT}\left(\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} \right)$$

Classical Confidence Intervals on $\mu_1 - \mu_2$ (not independent):

Paired if equal-size samples are pairs of observations on the *same set of individuals*.
Unpaired otherwise – but must be uncorrelated.

Classical Hypothesis Tests:

$\mathcal{H}_o : \mu = \mu_o$. Select the appropriate $\mathcal{H}_A$ and level of significance α.
The burden of proof is on $\mathcal{H}_A$.

Two tailed test, $\mathcal{H}_A : \mu \neq \mu_o$:

Method 1: Evaluate $c_L, c_U = \mu_o \pm t_{\alpha/2, n-1} \cdot \dfrac{s}{\sqrt{n}}$

Reject $\mathcal{H}_o$ iff $\bar{x} < c_L$ or $\bar{x} > c_U$

Method 2: Reject $\mathcal{H}_o$ iff

$$\left| t_{\text{obs}} \right| > t_{\alpha/2, n-1}, \quad \text{where} \quad t_{\text{obs}} = \frac{\bar{x} - \mu_\circ}{\left(\frac{s}{\sqrt{n}}\right)}$$

Method 3: Reject $\mathcal{H}_o$ iff $\text{P}\left[\left|T\right| > \left|t_{\text{obs}}\right|\right] < \alpha$, where $t_{\text{obs}} = \dfrac{\bar{x} - \mu_\circ}{\left(\frac{s}{\sqrt{n}}\right)}$

Upper tailed test, $\mathcal{H}_A : \mu > \mu_o$:

Method 1: Evaluate $c = \mu_o + t_{\alpha,\, n-1} \cdot \dfrac{s}{\sqrt{n}}$

Reject $\mathcal{H}_o$ iff $\bar{x} > c$

Method 2: Reject $\mathcal{H}_o$ iff $t_{\text{obs}} > t_{\alpha, n-1}$, where $t_{\text{obs}} = \dfrac{\bar{x} - \mu_\circ}{\left(\frac{s}{\sqrt{n}}\right)}$

Method 3: Reject $\mathcal{H}_o$ iff $\text{P}\left[T > t_{\text{obs}}\right] < \alpha$, where $t_{\text{obs}} = \dfrac{\bar{x} - \mu_\circ}{\left(\frac{s}{\sqrt{n}}\right)}$

Lower tailed test, $\mathcal{H}_\text{A} :\ \mu < \mu_\text{o}$:

Method 1: Evaluate $c \ = \ \mu_\circ \ - \ t_{\alpha, n-1} \cdot \dfrac{s}{\sqrt{n}}$

Reject $\mathcal{H}_\text{o}$ iff $\bar{x} < c$

Method 2: Reject $\mathcal{H}_\text{o}$ iff $t_\text{obs} \ < -t_{\alpha, n-1}$, where $t_\text{obs} \ = \dfrac{\bar{x} - \mu_\circ}{\left(\dfrac{s}{\sqrt{n}}\right)}$

Method 3: Reject $\mathcal{H}_\text{o}$ iff $\text{P}\left[T \ < \ t_\text{obs}\right] \ < \alpha$, where $t_\text{obs} \ = \dfrac{\bar{x} - \mu_\circ}{\left(\dfrac{s}{\sqrt{n}}\right)}$

Modify the above if the sample size is large and/or σ^2 is known.
Modify the above for two sample unpaired tests.

Chi-Square Goodness-of-Fit Test:

$\mathcal{H}_\text{o} :\ p_1 = p_{1\text{o}}, p_2 = p_{2\text{o}}, \ldots, p_k = p_{k\text{o}}$ vs. $\mathcal{H}_\text{A}$: not $\mathcal{H}_\text{o}$

Observed values $\{O_1, O_2, \ldots, O_k\}$. Calculate expected values $\{E_1, E_2, \ldots, E_k\}$, then compare $\chi^2 = \displaystyle\sum_{i=1}^{k} \frac{\left(O_i - E_i\right)^2}{E_i}$ to $c = \chi^2_{\alpha, k-1}$. Reject $\mathcal{H}_\text{o}$ iff $\chi^2 > c$

Chi-Square Test for Independence:

Observations $\{o_{ij}\}$ in I rows and J columns. $\mathcal{H}_\text{o}$: factors are independent.

Calculate $e_{ij} = \dfrac{o_{i\bullet} \times o_{\bullet j}}{o_{\bullet\bullet}}$. Iff $\chi^2 = \displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{\left(o_{ij} - e_{ij}\right)^2}{e_{ij}} > \chi^2_{\alpha, (I-1)(J-1)}$ then reject $\mathcal{H}_\text{o}$.

Simple Linear Regression:

Model: $Y \ = \ \beta_1 x + \beta_0 + \varepsilon$, $\varepsilon \sim \text{N}\left(0, \sigma^2\right)$

Check for linear trend with constant error variance and normally distributed residuals.
Find summary statistics:

$$n \qquad \sum x \qquad \sum y \qquad \sum x^2 \qquad \sum xy \qquad \sum y^2$$

Evaluate

$$n S_{xx} \ = n\sum x^2 - \left(\sum x\right)^2 \quad , \quad n S_{xy} \ = n\sum xy - \sum x \sum y \quad , \quad n S_{yy} \ = n\sum y^2 - \left(\sum y\right)^2$$

then

$$\hat{\beta}_1 \ = \ \frac{S_{xy}}{S_{xx}} \quad , \quad \hat{\beta}_0 \ = \ \frac{1}{n}\left(\sum y \ - \ \hat{\beta}_1 \sum x\right)$$

The line of best fit in the least squares sense is $y \ = \ \hat{\beta}_1 x \ + \ \hat{\beta}_0$

Entries in the ANOVA table:

$$SSR = \sum \left(\hat{y}-\bar{y}\right)^2 = \frac{\left(nS_{xy}\right)^2}{n\left(nS_{xx}\right)} \quad , \quad SST = \sum \left(y-\bar{y}\right)^2 = S_{yy} = \frac{\left(nS_{yy}\right)}{n}$$

$$SSE = \sum \left(y-\hat{y}\right)^2 = SST - SSR = \frac{\left(nS_{xx}\right)\left(nS_{yy}\right) - \left(nS_{xy}\right)^2}{n\left(nS_{xx}\right)}$$

$$MSR = \frac{SSR}{\nu_R} = SSR \ , \qquad s^2 = MSE = \frac{SSE}{\nu_E} = \frac{SSE}{n-2} \ , \qquad f = t^2 = \frac{MSR}{MSE}$$

Source	d.f.	Sum Sq.	Mean Sq.	f
R	1	SSR	MSR	f
E	$n-2$	SSE	MSE	
T	$n-1$	SST		

Variance of slope $\hat{\beta}_1$: $\quad {s_b}^2 = \dfrac{MSE}{S_{xx}} = \dfrac{\left(nS_{xx}\right)\left(nS_{yy}\right)-\left(nS_{xy}\right)^2}{(n-2)\left(nS_{xx}\right)^2}$

Coefficient of Determination:

$$r^2 = \frac{SSR}{SST} = \frac{\left(nS_{xy}\right)^2}{\left(nS_{xx}\right)\left(nS_{yy}\right)}$$

Correlation coefficient $= r = \text{sign}\left(\hat{\beta}_1\right)\cdot\sqrt{r^2}$

Hypothesis tests on the linear association of X and Y (all have $\nu = \nu_\text{E} = n-2$):

To test $\mathcal{H}_\text{o}: \rho = 0$ vs. $\mathcal{H}_\text{A}: \rho \neq 0$ (or, equivalently, $\mathcal{H}_\text{o}: \beta_1 = 0$ vs. $\mathcal{H}_\text{A}: \beta_1 \neq 0$):
Use *any* of

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \ , \quad t = \frac{\hat{\beta}_1 - 0}{s_b} \ , \quad t = \sqrt{\frac{MSR}{MSE}} \ , \quad t = \sqrt{\frac{\left(nS_{xy}\right)^2(n-2)}{\left(nS_{xx}\right)\left(nS_{yy}\right)-\left(nS_{xy}\right)^2}}$$

To test $\mathcal{H}_\text{o}: \beta_1 = \beta_{1\text{o}}$ vs. $\mathcal{H}_\text{A}: \beta_1 > \beta_{1\text{o}}$ (when $\beta_{1\text{o}} \neq 0$) use

$$t = \frac{\hat{\beta}_1 - \beta_{1\text{o}}}{s_b} \qquad \text{or} \qquad t = \left(\left(nS_{xy}\right) - \beta_{1\text{o}}\left(nS_{xx}\right)\right)\sqrt{\frac{(n-2)}{\left(nS_{xx}\right)\left(nS_{yy}\right)-\left(nS_{xy}\right)^2}}$$

The $(1-\alpha)\times 100\%$ **confidence interval** estimate for $\mu = \text{E}\left[Y \mid x = x_\text{o}\right]$ is

$$\left(\hat{\beta}_0 + \hat{\beta}_1 x_\text{o}\right) \pm t_{\alpha/2,\,(n-2)}\, s\sqrt{\frac{1}{n} + \frac{n\left(x_\text{o}-\bar{x}\right)^2}{\left(nS_{xx}\right)}}$$

The $(1-\alpha)\times 100\%$ **prediction interval** estimate for $Y \mid x = x_\text{o}$ is

$$\left(\hat{\beta}_0 + \hat{\beta}_1 x_\text{o}\right) \pm t_{\alpha/2,\,(n-2)}\, s\sqrt{1 + \frac{1}{n} + \frac{n\left(x_\text{o}-\bar{x}\right)^2}{\left(nS_{xx}\right)}}$$

[Space for Additional Notes]