STAT 251 Formula Sheet

Measures of Center

Mean: $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1+x_2+...+x_n}{n}$

Median: If n is even then $\tilde{x} = \frac{(\frac{n}{2})^{\text{th}} \text{ obs.} + (\frac{n+1}{2})^{\text{th}} \text{ obs.}}{2}$

If n is odd then $\tilde{x} = \frac{n+1}{2}^{\text{th}}$ obs.

Measures of Variability

Range: $R = x_{\text{largest}} - x_{\text{smallest}}$

Variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$

Standard deviation: $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}}$

IQR: $\text{IQR} = \text{Q3} - \text{Q1}$

Method to compute $\mathbf{Q}_{(p)}$:

- Sort data from smallest to largest: $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$
- Compute the number $np + 0.5$
- If $np + 0.5$ is an integer, m, then: $\text{Q}_{(p)} = x_{(m)}$
- If $np+0.5$ is not an integer, $m < np+0.5 < m+1$ for some integer m, then: $\text{Q}_{(p)} = \frac{x_{(m)}+x_{(m+1)}}{2}$

Outliers:

- Values smaller than $\text{Q1} - (1.5 \times \text{IQR})$ are outliers
- Values greater than $\text{Q3} + (1.5 \times \text{IQR})$ are outliers

Discrete Random Variables

Consider a **discrete** random variable X

Probability Mass Function (pmf): $f(x) = P(X = x)$

1. $f(x) \geq 0$ for all x in X
2. $\sum_x f(x) = 1$

Cumulative Distributive Function (cdf): $F(x) = P(X \leq x) = \sum_{k \leq x} f(k)$

Mean (μ): $E(X) = \sum_x x f(x)$

Expected value: $E(g(X)) = \sum_x g(x) f(x)$

Variance (σ^2): $Var(X) = \sum_x (x-\mu)^2 f(x) = E(X^2) - [E(X)]^2$

SD (σ): $SD(X) = \sqrt{Var(X)}$

Sets and Probability

Properties of Probability:

- *General Addition Rule:* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- *Complement Rule:* $P(A^c) = 1 - P(A)$
- If $A \subseteq B$ then $P(A \cap B) = P(A)$
- If $A \subseteq B$ then $P(A) \leq P(B)$
- $P(\emptyset) = 0$ and $P(S) = 1$
- $0 \leq P(A) \leq 1$ for all A

Conditional Probability:

- $P(A|B) = \frac{P(A\cap B)}{P(B)}$ and $P(B|A) = \frac{P(A \cap B)}{P(A)}$
- *Multiplication Rule:* $P(A \cap B) = P(B) \times P(A|B)$ and $P(A \cap B) = P(A) \times P(B|A)$
- Events A and B are **independent** if and only if $P(A \cap B) = P(A)P(B)$ and thus $P(A|B) = P(A)$ and $P(B|A) = P(B)$

Bayes' Theorem:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+...+P(B|A_n)P(A_n)}$$

Continuous Random Variables

Consider a **continuous** random variable X

Probability Density Function (pdf): $P(a \leq X \leq b) = \int_a^b f(x)dx$

1. $f(x) \geq 0$ for all x
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Cumulative Distributive Function (cdf): $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$

Median: x such that $F(x) = 0.5$

Q_1 and Q_3: x such that $F(x) = 0.25$ and x such that $F(x) = 0.75$

Mean (μ): $E(X) = \int_{-\infty}^{\infty} x f(x)dx$

Expected value: $E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$

Variance (σ^2): $Var(X) = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx = E(X^2) - [E(X)]^2$

SD (σ): $SD(X) = \sqrt{Var(X)}$

Summarizing Main Features of $f(x)$

Consider two random variables X, Y

Properties of Probability:

- $E(aX+b) = aE(X)+b$, for $a, b \in \mathbb{R}$
- $E(X+Y) = E(X)+E(Y)$, for all pairs of X and Y
- $E(XY) = E(X)E(Y)$, for independent X and Y
- $Var(aX+b) = a^2Var(X)$, for $a, b \in \mathbb{R}$
- $Var(X+Y) = Var(X)+Var(Y)$
  $Var(X-Y) = Var(X)+Var(Y)$, for independent X and Y

Covariance:

- $\mathrm{Cov}(X,Y) = E[X-E(X)][Y-E(Y)] = E(XY)-E(X)E(Y)$
  If X and Y are independent, $Cov(X,Y) = 0$
- $Var(X+Y) = Var(X)+Var(Y)+2Cov(X,Y)$
- $Var(aX+bY+c) = a^2Var(X)+b^2Var(Y)+2abCov(X,Y)$

Sum and Average of Independent Random Variables

Sum of Independent Random Variables:

$Y = a_1X_1+a_2X_2+...+a_nX_n$, for $a_1, a_2, ..., a_n \in \mathbb{R}$

- $E(Y) = a_1E(X_1)+a_2E(X_2)+...+a_nE(X_n)$
- $Var(Y) = a_1^2Var(X_1)+a_2^2Var(X_2)+...+a_n^2Var(X_n)$

If n random variables X_i have common mean μ and common variance σ^2 then,

- $E(Y) = (a_1+a_2+...+a_n)\mu$
- $Var(Y) = (a_1^2+a_2^2+...+a_n^2)\sigma^2$

Average of Independent Random Variables:

$X_1, X_2, ..., X_n$ are n independent random variables

- $\overline{X} = \frac{X_1+X_2+...+X_n}{n}$
- $E[\overline{X}] = \frac{1}{n}[E(X_1)+E(X_2)+...+E(X_n)]$
- $Var[\overline{X}] = \frac{1}{n^2}[Var(X_1)+Var(X_2)+...+Var(X_n)]$

If n random variables X_i have common mean μ and common variance σ^2 then,

- $E[\overline{X}] = \mu$
- $Var[\overline{X}] = \frac{\sigma^2}{n}$

Maximum and Minimum of Independent Variables

Given n independent random variables $X_1, X_2, ..., X_n$.
For each X_i, cdf $F_X(x)$ and pdf is $f_X(x)$.

Maximum of Independent Random Variables:

Consider $V = max\{X_1, X_2, ..., X_n\}$

cdf of V

$$\begin{aligned} F_V(v) &= P(V \le v) = P(X_1 \le v, X_2 \le v, ..., X_n \le v) \\ &= P(X_1 \le v)P(X_2 \le v)...P(X_n \le v) = F_{X_1}(v)F_{X_2}(v)...F_{X_n}(v) \\ &= [F_X(v)]^n \text{ ; if } X_i\text{'s are all identically distributed} \end{aligned}$$

pdf of V

$$\begin{aligned} f_V(v) &= F_V'(v) = \tfrac{d}{dv}F_V(v) = \tfrac{d}{dv}[F_X(v)]^n = n[F_X(v)]^{n-1}\tfrac{d}{dv}F_X(v) \\ &= n[F_X(v)]^{n-1}f_X(v) \end{aligned}$$

Minimum of Independent Random Variables:

Consider $U = min\{X_1, X_2, ..., X_n\}$

cdf of U

$$\begin{aligned} F_U(u) &= P(U \le u) = 1-P(U > u) = 1-P(X_1 > u, X_2 > u, ..., X_n > u) \\ &= 1-P(X_1 > u)P(X_2 > u)...P(X_n > u) \\ &= 1-[1-F_{X_1}(u)][1-F_{X_2}(u)]...[1-F_{X_n}(u)] \\ &= 1-[1-F_X(u)]^n \text{ ; if } X_i\text{'s are all identically distributed} \end{aligned}$$

pdf of U

$$\begin{aligned} f_U(u) &= F_U'(u) = \tfrac{d}{du}\{1-[1-F_X(u)]^2\} = 0-n[1-F_X(u)]^{n-1}\tfrac{d}{du}(-F_X(u)) \\ &= n[1-F_X(u)]^{n-1}f_X(u) \end{aligned}$$

Some Continuous Distributions

Uniform Distribution: $X \sim U(a,b)$

Mean: $\mu = E(X) = \frac{a+b}{2}$

Variance: $\sigma^2 = Var(X) = \frac{(b-a)^2}{12}$

pdf of X

$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

cdf of X

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$

Exponential Distribution: $X \sim Exp(\lambda)$

Mean: $\mu = E(X) = \frac{1}{\lambda}$

Variance: $\sigma^2 = Var(X) = \frac{1}{\lambda^2}$

pdf of X

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

cdf of X

$$F(x) = \begin{cases} 1-e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

Normal Distribution

Normal Distribution: $X \sim N(\mu, \sigma^2)$
Standardized Normal: $Z \sim N(0,1)$ where $Z = \frac{X-\mu}{\sigma}$

68-95-99.7 Rule:

- approximately 68% of observations fall within σ of μ
- approximately 95% of observations fall within 2σ of μ
- approximately 99.7% of observations fall within 3σ of μ

Bernoulli and Binomial Random Variables

Bernoulli Random Variable:
Bernoulli random variable X has only two outcomes, success and failure.
$P(Success) = p$ and $P(Failure) = 1 - p$

Bernoulli Distribution:
$X \sim Bernoulli(p)$
pmf: $P(X = x) = p^x(1-p)^{1-x}$ for $x = 0, 1$

Mean: $\mu = E(X) = p$
Variance: $\sigma^2 = Var(X) = p(1-p)$

Binomial Random Variable:
Binomial random variable X is the number of successes for n independent trials and each trial has the same probability of success p.

Binomial Distribution:
$X \sim Bin(n, p)$
pmf: $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, 2, ..., n$
cdf: $P(X \leq x) = \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i}$ for $x = 0, 1, 2, ..., n$
Note: $\binom{n}{x} = \frac{n!}{x!(n-x)!}$

Mean: $\mu = E(X) = np$
Variance: $\sigma^2 = Var(X) = np(1-p)$

Geometric Distribution

Geometric Random Variable:
Geometric random variable X is the number of independent trials needed until the first success occurs.

Geometric Distribution:
$X \sim Geo(p)$ where p is the probability of success
pmf: $P(X = x) = p(1-p)^{x-1}$ for $x = 1, 2, 3, ...$
cdf: $P(X \leq x) = 1 - (1-p)^x$ for $x = 1, 2, 3, ...$

Mean: $\mu = E(X) = \frac{1}{p}$
Variance: $\sigma^2 = Var(X) = \frac{1-p}{p^2}$

Poisson Distribution

Poisson Process:
Random variable X is the number of occurrences in a given interval.

Poisson Distribution:
$X \sim Poisson(\lambda)$ where λ is the rate of occurrences
pmf: $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$ for $x = 0, 1, 2, 3, ...$
cdf: $P(X \leq x) = \sum_{i=0}^{x} \frac{\lambda^i e^{-\lambda}}{i!}$ for $x = 0, 1, 2, 3, ...$

Mean: $\mu = E(X) = \lambda$
Variance: $\sigma^2 = Var(X) = \lambda$

- Let $T \sim Exp(\lambda)$ be the time between two consecutive occurrences of events. (Can also be the waiting time for first event.)

Poisson Approximation to the Binomial Distribution

Let $X \sim Bin(n, p)$ be a binomial random variable. If n is large ($n \geq 20$) and p or $1 - p$ is small ($np < 5$ or $n(1-p) < 5$), then we can use a Poisson random variable with rate $\lambda = np$ to approximate the probabilistic behaviour of X.

$X \sim Poisson(np)$, approx. for $x = 0, 1, 2, ...n$

Central Limit Theorem

Let X_1, X_2,..., X_n be a random sample from an arbitrary population/distribution with mean μ and variance σ^2. When n is large ($n \geq 20$) then

$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n} \sim N(\mu, \frac{\sigma^2}{n})$, approx.

When dealing with sum, the CLT can still be used. Then

$T = X_1 + X_2 + ... + X_n = n\overline{X}$
$T \sim N(n\mu, n\sigma^2)$, approx.

Normal Approximation to the Binomial Distribution

Let $X \sim Bin(n, p)$. When n is large so that both $np \geq 5$ and $n(1-p) \geq 5$. We can use the normal distribution to get an approximate answer. Remember to use **continuity correction**.

$X \sim N(np, np(1-p))$, approx.

Normal Approximation to the Poisson Distribution

Let $X \sim Poisson(\lambda)$. When λ is large ($\lambda \geq 20$) then the Normal distribution can be used to approximate the Poisson distribution. Remember to use **continuity correction**.

$X \sim N(\lambda, \lambda)$, approx.

Continuity Correction

Consider continuous random variable Y and discrete random variable X.

- $P(X > 4) = P(X \geq 5) = P(Y \geq 4.5)$
- $P(X \geq 4) = P(Y \geq 3.5)$
- $P(X < 4) = P(X \leq 3) = P(Y \leq 3.5)$
- $P(X \leq 4) = P(Y \leq 4.5)$
- $P(X = 4) = P(3.5 \leq Y \leq 4.5)$

Point Estimators

Suppose that X_1, X_2,..., X_n are random samples from a population with mean μ and variance σ^2.

- $\overline{x}$ is an unbiased estimator of μ
  $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
- s^2 is an unbiased estimator of σ^2
  $s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$
- θ is the parameter, $\hat{\theta}$ is the point estimator. When $E(\hat{\theta}) = \theta$, $\hat{\theta}$ is an unbiased estimator. The bias of an estimator is $bias(\theta) = E(\hat{\theta}) - \theta$.

Confidence Interval

$(1-\alpha)100\%$ **Confidence Interval for population mean μ:**
(point estimator of μ is $\overline{x}$)

General Form: point estimate $\pm$ margin of error
When σ^2 is **known**: $\overline{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
When σ^2 is **unknown**: $\overline{x} \pm t_{\frac{\alpha}{2},n-1} \frac{s}{\sqrt{n}}$

Typical z values of α:

$\alpha = 0.1$	90%	$z_{\frac{\alpha}{2}} = z_{0.05} = 1.645$
$\alpha = 0.05$	95%	$z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$
$\alpha = 0.01$	99%	$z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

$(1-\alpha)100\%$ **Confidence Interval for $\mu_1 - \mu_2$:**
$(\overline{x}_1 - \overline{x}_2) \pm t_{\frac{\alpha}{2},n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

Pooled Standard Deviation

Requires assumptions that population variances are equal: $\sigma_1^2 = \sigma_2^2 = \sigma^2$
The pooled standard deviation s_p estimates the common standard deviation σ.
$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

Testing of Hypotheses about μ

H_o: Null hypothesis is a tentative assumption about a population parameter.
H_a: Alternative hypothesis is what the test is attempting to establish.

- $H_o : \mu \geq \mu_o$ vs $H_a : \mu < \mu_o$ (one-tail test, lower-tail)
- $H_o : \mu \leq \mu_o$ vs $H_a : \mu > \mu_o$ (one-tail test, upper-tail)
- $H_o : \mu = \mu_o$ vs $H_a : \mu \neq \mu_o$ (two-tail test)

Test Statistic:
Case 1: σ^2 is <u>known</u>
$z = \frac{\overline{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

Case 2: σ^2 is <u>unknown</u>
$t = \frac{\overline{x} - \mu_o}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$

Type I and Type II errors:

Type I error:	rejecting H_o when H_o is true
Type II error:	not rejecting H_o with H_o is false

$P(\text{Type I error}) = \alpha$
$P(\text{Type II error}) = \beta$

Power is the probability of rejecting H_o, when H_o is false.
Power $= 1 - \beta$

Comparison of two means:
Two independent populations with means μ_1 and μ_2.
Assume random samples, normal distributions, and equal variances ($\sigma_1^2 = \sigma_2^2$).

- $H_o : \mu_1 - \mu_2 \geq \Delta_o$ vs $H_a : \mu_1 - \mu_2 < \Delta_o$ (lower-tail)
- $H_o : \mu_1 - \mu_2 \leq \Delta_o$ vs $H_a : \mu_1 - \mu_2 > \Delta_o$ (upper-tail)
- $H_o : \mu_1 - \mu_2 = \Delta_o$ vs $H_a : \mu_1 - \mu_2 \neq \Delta_o$ (two-tail)

Test Statistic:
$t = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_o}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$
s_p is the pooled standard deviation

Rejection Rules:
Consider test statistic z, and significance value α.

- Lower-tail test: Reject H_o if $z \leq z_\alpha$
- Upper-tail test: Reject H_o if $z \geq z_\alpha$
- Two-tail test: Reject H_o if $|z| \geq z_{\frac{\alpha}{2}}$

Analysis of Variance (ANOVA)

One-way ANOVA:
k = number of populations or treatments being compared
μ_1 = mean of population 1 or true average response when treatment 1 is applied.
...
μ_k = mean of population k or true average response when treatment k is applied.

Assumptions:

- For each population, response variable is normally distributed
- Variance of response variable, σ^2 is the same for all the populations
- The observations must be independent

Hypotheses:
$H_o : \mu_1 = \mu_2 = ... = \mu_k$
$H_a : \mu_i \neq \mu_j$ for $i \neq j$

Notation:
y_{ij} is the j^{th} observed value from the i^{th} population/treatment.

Total mean: $\overline{y}_{i\cdot} = \frac{y_{i\cdot}}{n_i} = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$
Total sample size: $n = n_1 + n_2 + ... + n_k$
Grand total: $y_{\cdot\cdot} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}$
Grand mean: $\overline{y}_{\cdot\cdot} = \frac{y_{\cdot\cdot}}{n} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}}{n}$

$s^2 = \frac{\sum_{j=1}^{k}(n_i-1)s_i^2}{n-k}$ = MSE, where $s_i^2 = \frac{\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_{i\cdot})^2}{n_i-1}$

The ANOVA Table:

Source of Variation	df	Sum of Squares	Mean Square	F-ratio
Treatment	$k-1$	SSTr	MSTr $= \frac{\text{SSTr}}{k-1}$	$\frac{\text{MSTr}}{\text{MSE}}$
Error	$n-k$	SSE	MSE $= \frac{\text{SSE}}{n-k}$	
Total	$n-1$	SST		

$$\text{SST} = \text{SSTr} + \text{SSE}$$

$$\text{SST} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{\cdot\cdot})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - \frac{1}{n} y_{\cdot\cdot}^2$$
$$\text{SSTr} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot})^2 = \sum_{i=1}^{k} \frac{1}{n_i} y_{i\cdot}^2 - \frac{1}{n} y_{\cdot\cdot}^2$$
$$\text{SSE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{i\cdot})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{k} \frac{y_{i\cdot}^2}{n_i} = \sum_{i=1}^{k}(n_i - 1)s_i^2$$

Test Statistic:
$F_{obs} = \frac{\text{MSTr}}{\text{MSE}} \sim F_{v_1, v_2}$ $\quad v1 = df(\text{SSTr}) = k-1$
$\quad v2 = df(\text{SSE}) = n-k$

Reject H_o if $F_{obs} \geq F_{\alpha, v_1, v_2}$

Covariance and Correlation Coefficient

On a scatter plot, each observation is represented as a point with x-coord x_i and y-coord y_i.

Sample Covariance: $Cov(x,y)$

$$\begin{aligned} Cov(x,y) &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) \\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}\right] \\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}\right] \end{aligned}$$

- If x and y are positively associated, then $Cov(x,y)$ will be large and positive
- If x and y are negatively associated, then $Cov(x,y)$ will be large and negative
- If the variables are not positively nor negatively associated, then $Cov(x,y)$ will be small

Sample Correlation Coefficient: r

$r = \frac{1}{n-1}\sum_{i=1}^{n}(\frac{x_i - \overline{x}}{s_x})(\frac{y_i - \overline{y}}{s_y})$, where $s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}}$ and $s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\overline{y})^2}{n-1}}$
$r = \frac{Cov(x,y)}{s_x s_y}$

- Always falls between -1 and +1
- A positive r value indicates a positive association
- A negative r value indicates a negative association
- r value close to +1 or -1 indicates a strong linear association
- r value close to 0 indicates a weak association

Simple Linear Regression

Regression Line:
Simple linear regression model: $y = \beta_o + \beta_1 x + \varepsilon$
β_o, β_1, and σ^2 are parameters, y and ε are random variables. ε is the error term.

True regression line: $E(y) = \beta_o + \beta_1 x$
Least squares regression line: $\hat{y} = \hat{\beta}_o + \hat{\beta}_1 x$
$\hat{y}$, $\hat{\beta}_o$, and $\hat{\beta}_1$ are point estimates for y, β_o, and β_1.
Residual: $\varepsilon_i = y_i - \hat{y}_i$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2} = r\frac{s_y}{s_x}$$
$$\hat{\beta}_o = \frac{\sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i}{n} = \overline{y} - \hat{\beta}_1 \overline{x}$$

Coefficient of Determination: r^2
The proportion of observed y variation that can be explained by the simple linear regression model.

Estimating σ^2 (SLR)

$\hat{\sigma}^2 = s^2 = \frac{\text{SSE}}{n-2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{n-2}$

Error Sum of Squares (SSE):
$\text{SSE} = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}[y_i - (\hat{\beta}_o + \hat{\beta}_1 x_i)]^2$
SSE is a measure of variation in y left unexplained by linear regression model.

Total Sum of Squares (SST):
$\text{SST} = \sum_{i=1}^{n}(y_i - \overline{y})^2$
SST is sum of squared deviations about sample mean of observed y values.

Regression Sum of Squares (SSR):
$\text{SSR} = \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$
SSR is total variation explained by the linear regression model.

$\text{SST} = \text{SSR} + \text{SSE}$

Coefficient of Determination from SST, SSR, and SSE:
$r^2 = 1 - \frac{SSE}{SST}$
or
$r^2 = \frac{SSR}{SST}$

Slope Parameter β_1 (SLR)

When $\beta_1 = 0$ there is no linear relationship between the two variables.

Hypotheses:
$H_o : \beta_1 = 0$
$H_a : \beta_1 \neq 0$

Test statistic:
$t_{\text{obs}} = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}} \sim t_{n-2}$, where $s_{\hat{\beta}_1} = \frac{s}{s_x\sqrt{n-1}}$

Reject H_o if $|t_{\text{obs}}| \geq t_{\frac{\alpha}{2},n-2}$

$(1-\alpha)100\%$ **Confidence Interval for β_1:**
$\hat{\beta}_1 \pm t_{\frac{\alpha}{2},n-2} s_{\hat{\beta}_1}$