ICGN103: Formulas / Final Examination

1. $\mu = \bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

2. $s^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \dfrac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{n-1} = \dfrac{\sum_{i=1}^{n} x_i^2 - \dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$

3. $s = \sqrt{s^2}$

4. ***Empirical Rule for a Normally Distributed Population***
   If a population has mean μ and standard deviation σ and is described by a normal curve, then approximately 68.26%, 95.44%, and 99.73% of the population measurements lie in the interval $[\mu \pm \sigma]$, $[\mu \pm 2\sigma]$ and $[\mu \pm 3\sigma]$, respectively.

5. $z = \dfrac{x - \text{mean}}{\text{standard deviation}}$

6. The mean of $\bar{x}$, $\mu_{\bar{x}} = \mu$,
   The standard deviation of $\bar{x}$, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

7. $\hat{p} = \frac{x}{n}$

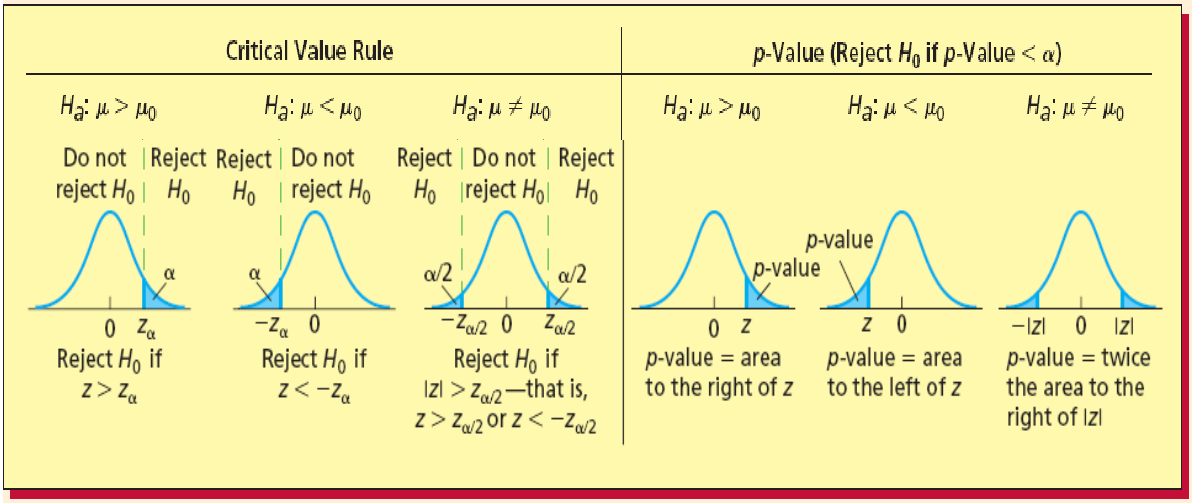
8. The mean of $\hat{p}$, $\mu_{\hat{p}} = p$,
   The standard deviation of $\hat{p}$, $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$

9. Confidence Interval
   - $\bar{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} = \left[\bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$
   - $\left[\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\right]$

10. Sample size
    - $n = \left(\dfrac{z_{\alpha/2}\sigma}{E}\right)^2$
    - $n = p(1-p)\left(\dfrac{z_{\alpha/2}}{E}\right)^2$

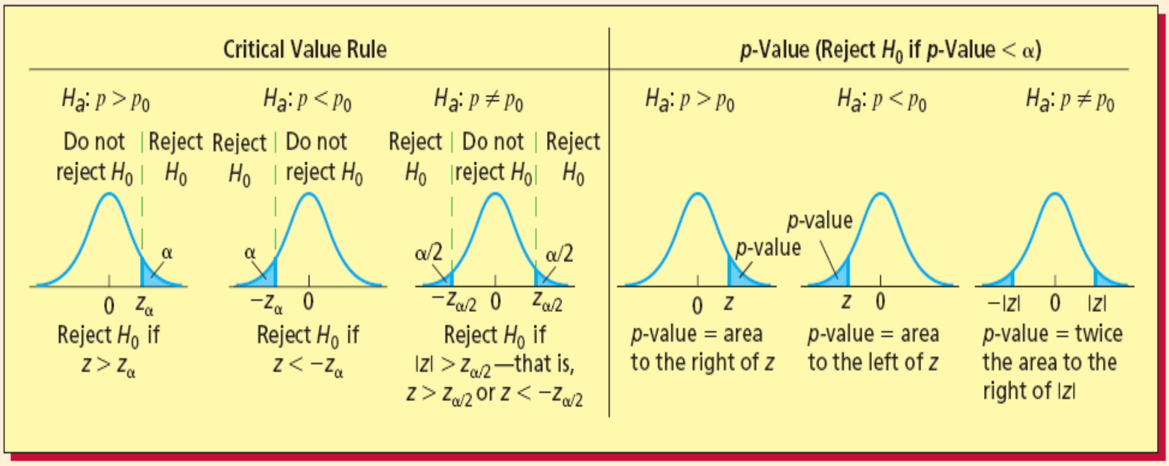
11. Testing a Hypothesis about a Population Mean When σ is Known

Null Hypothesis	Test Statistic	Assumptions
$H_0: \mu = \mu_0$	$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	Normal population or Large sample size



12. Testing a Hypothesis about a Population Proportion

Null Hypothesis	Test Statistic	Assumptions[3]
$H_0: p = p_0$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	$np_0 \geq 5$ and $n(1-p_0) \geq 5$



13. The least squares regression line: $\hat{y} = b_0 + b_1 x$

14. The least squares point estimate of the slope β_1

$$b_1 = \frac{SS_{xy}}{SS_{xx}} \quad \textit{where}$$

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n}$$

$$SS_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}$$

15. The least squares point estimate of the y-intercept β_0

$$b_0 = \bar{y} - b_1\bar{x} \quad where$$
$$\bar{y} = \frac{\sum y_i}{n}$$
$$\bar{x} = \frac{\sum x_i}{n}$$

16. Residual/Error Term

Sum Square Error $\quad SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

Mean Square Error $\quad s^2 = MSE = \frac{SSE}{n\text{-}2}$

Standard Error $\quad s = \sqrt{MSE} = \sqrt{\frac{SSE}{n\text{-}2}}$

17. The Simple Coefficient of Determination: $\quad r^2 = \frac{\textbf{Explained variation}}{\textbf{Total variation}}$

Where $\quad$ Total variation $= \sum(y_i - \bar{y})^2$

Explained variation $= \sum(\hat{y}_i - \bar{y})^2$

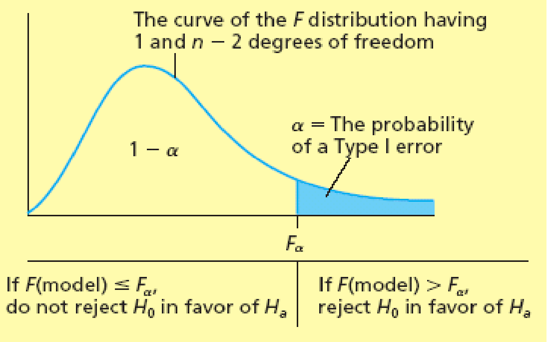
Unexplained variation or $SSE = \sum(y_i - \hat{y}_i)^2$

Total variation = Explained variation + Unexplained variation

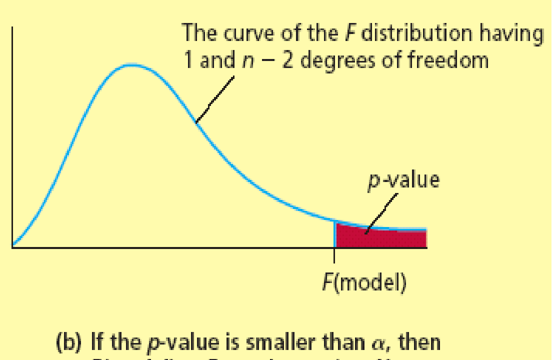
18. The simple correlation coefficient: $\quad r = +\sqrt{r^2}$ or $r = -\sqrt{r^2}$

19. An F-test for the Simple Linear Regression Model:

$$F(\text{model}) = \frac{\textbf{Explained variation}}{(\textbf{Unexplained variation})/(\boldsymbol{n-2})}$$



(a) The rejection point F_α based on setting the probability of a Type I error equal to α



(b) If the p-value is smaller than α, then $F(\text{model}) > F_\alpha$ and we reject H_0.