Formula Sheet III

Chapter 6: Sampling Distributions

Sample Means

Mean and Standard Distribution of Sample Means

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$E(\bar{X}) = \mu$$

The Central Limit Theorem states that for *n* sufficiently large (at least 30 in practice), the sampling distribution of the means will be approximately normal.

Sample Proportions

Sampling distribution of the sample proportion:

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$E(\bar{p}) = p$$

The distribution will be approximately normal for *n* large enough such that 3 standard deviations above and below the mean lies in the interval (0,1).

Chapter 7: Confidence Interval Estimation

Confidence Intervals for the Population Mean

Confidence Interval	σ is known	σ is not known
Large sample (n>30)	$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}$
Finite population with large sample (n>30, n/N>5%)	$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N}}$	$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N}}$
Sample is small, population approximately normal	$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

For the above formulae, $1-\alpha$ is the level of confidence and the term to the right of $\pm$ is the margin of error. When *t* is used, the degrees of freedom are *n*–1.

Confidence Intervals for Population Proportions

$$\hat{p} \pm Z_{\alpha/2} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

n must be sufficiently large such that the interval for $Z_{\alpha/2} = 3$ lies in the interval (0,1). Otherwise, if *n* is small, then the following adjustment is used:

$$\hat{p} \pm Z_{\alpha/2} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \qquad \hat{p} = \frac{x+2}{n+4}$$

If the population is finite and *n*/*N*>.05, then the following adjustment is used:

$$\hat{p} \pm Z_{\alpha/2} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N}}$$

Chapter 8: Hypothesis Testing

Regions of rejection

This is the same for means and proportions, just replace μ with p

Hypothesis	$H_0: \mu = 0$ $H_A: \mu \neq 0$	$H_0: \mu = 0$ $H_A: \mu < 0$	$H_0: \mu = 0$ $H_A: \mu > 0$
Region for Rejection (large samples)	$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$	$z < -z_{\alpha}$	$z > z_{\alpha}$
Region for Rejection (small samples, σ unknown)	$t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$	$t < -t_{\alpha}$	$t > t_{\alpha}$

Large Samples

- Population Mean

$$z = \frac{\overline{x} - \mu}{s / \sqrt{n}}$$

- Population Proportion

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Small Samples using the t-statistic

- Population Mean

$$t = \frac{\overline{x} - \mu}{s / \sqrt{n}}$$

For a desired level of SE, the sample size is:

$$n = \frac{\left(Z_{\alpha/2}\right)^2 \sigma^2}{(SE)^2}$$

$$n = \frac{\left(Z_{\alpha/2}\right)^2 (p(1-p))}{\left(SE\right)^2}$$

If σ is unknown, we can estimate it using *s* or for a more conservative estimate use 1/4th the range of values. Since p is unknown, we can estimate $\left(p(1-p)\right)$ using $\left(\hat{p}(1-\hat{p})\right)$ or for a more conservative estimate use $p = .5$.

A *p*-value is the probability of getting a particular test statistic value or more extreme. It is the lowest significance with which the null can be rejected. Specifically, given a value of the test statistic *Z**, then the *p*-value is:

One Tailed: $p\text{-value} = \Pr(|Z| > |Z^*|)$

Two Tailed: $p\text{-value} = 2 \times \Pr(|Z| > |Z^*|)$