REFERENCE PAGES

STATISTICS

DESCRIPTIVE STATISTICS

1. Sample mean: $\overline{X} = \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right)$

2. Sample variance: $s_n^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right)$

3. Skewness (asymmetry): $\gamma_1 = \frac{\mu^3}{\sigma^3}$

4. Kurtosis (peakedness): $\gamma_2 = \frac{\mu^4}{\sigma^4}$

RULES OF THUMB

$P(|X-\mu| < 1\sigma) = 0.6826$

5. $P(|X-\mu| < 2\sigma) = 0.9544$

$P(|X-\mu| < 3\sigma) = 0.9973$, "$3\sigma$ rule"

CONFIDENCE INTERVAL ESTIMATION

6. Confidence interval of μ when σ^2 is known, using normal distribution:

a) 95% confidence interval: $\left[\overline{X} - 1.96\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}\right]$,

b) 90% confidence interval: $\left[\overline{X} - 1.645\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.645\frac{\sigma}{\sqrt{n}}\right]$.

7. Confidence interval of μ when σ^2 is unknown, using t-distribution:

$$T = \frac{\overline{X} - \mu}{\sqrt{s_n^2/n}} \sim t(n-1).$$

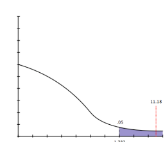
HYPOTHESIS TESTING

8. One-tail Procedure Example:

1) $H_0 : \mu = 0,\ H_1 : \mu > 0$ [one tail]

2) $\overline{X} = 0.031,\ s_n^2 = 0.0001,\ n = 13$ [b/c σ^2 is unknown, use t-distribution]

$$t = \frac{\overline{X} - \mu}{\sqrt{s_n^2/n}} = \frac{(0.031-0)\sqrt{13}}{.01} = 11.18 \sim t(12)$$



3) | $\alpha = 0.05,\ t_\alpha = 1.782 < 11.18 = t$

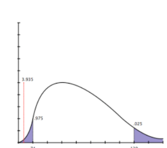
4) We reject H_0.

9. Two-tail Procedure Example:

1) $H_0 : \sigma^2 = 0.4,\ H_1 : \sigma^2 \neq 0.4$ [two tail]

2) $s_n^2 = 0.0159,\ n = 100$ [b/c testing for population variance, use χ^2-distribution]

$$\chi^2 = \frac{(n-1)s_n^2}{\sigma^2} \sim \chi^2(n-1) = \frac{(99)(0.0159)}{0.4} = 3.935 \sim \chi^2(99)$$



3) | $\alpha = 0.05,\ \ 3.935 < 74$

4) We reject H_0.

REFERENCE PAGES

S T A T I S T I C S

NORMAL DISTRIBUTION OR $N > 30$ AND σ_i KNOWN

10. One-sample z-test: $z = \dfrac{\bar{X} - \mu_0}{\left(\sigma/\sqrt{n}\right)}$, distribution of sample mean.

11. Two-sample z-test: $z = \dfrac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$, distribution of independent sample means.

t DISTRIBUTION OR $N < 30$ AND σ_i UNKNOWN, BUT EQUAL

12. One-sample t-test: $T = \dfrac{\bar{X} - \mu_0}{\left(s_n/\sqrt{n}\right)}$, $df = n - 1$, distribution of sample mean.

13. If $X \sim N(0,1)$ and $Y \sim \chi^2(n)$, then $T = \dfrac{X}{\sqrt{Y/n}}$, $df = n$, distribution of X related to sample mean and Y related to sample variance, then their ratio has to do with t-distribution.

14. Two-sample t-test: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}\sqrt{\frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x + n_y - 2}}} \sim t\left(n_x + n_y - 2\right)$, assuming $X \sim N\left(\mu_1, \sigma^2\right)$, $Y \sim N\left(\mu_2, \sigma^2\right)$, and $H_0 : \mu_1 = \mu_2$.

15. As $n \to 30$, t-distribution converges to normal distribution.

χ^2 DISTRIBUTION

16. One-sample χ^2 test: $\chi^2 = \dfrac{(n-1)s_n^2}{\sigma^2} \sim \chi^2(n-1)$, distribution of sample variance.

F-DISTRIBUTION

17. If $X \sim \chi^2(m)$ and $Y \sim \chi^2(n)$, then $F = \dfrac{(X/m)}{(Y/n)} \sim F(m,n)$, ratios of two sets of sample variances have to do with F-distribution.

18. ANOVA uses F-distribution.