Organizing Data

Class Width

$$\frac{highest\ value\ -\ lowest\ value}{number\ of\ classes}$$

Relative Frequency

$$\frac{f}{n} = \frac{frequency}{sample\ size}$$

Numerical Summaries of Data

Sample Mean

$$\bar{x} = \frac{\sum x_i}{n}$$

Weighted Mean

$$\bar{x}_w = \frac{\sum (x_i \cdot w_i)}{\sum w_i}$$

Population Mean

$$\mu = \frac{\sum x_i}{N}$$

Sample Variance

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

Population Variance

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N}$$

Sample Standard Deviation

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

Population Standard Deviation

$$\sigma = \sqrt{\sigma^2}$$

Range

$$max - min$$

Interquartile Range

$$IQR = Q_3 - Q_1$$

Five Number Summary

$$minimum, Q_1, median, Q_3, maximum$$

Elementary Probability

Probability of the Complement

$$P(A^c) = 1 - P(A)$$

General Multiplication Rule

$$P(A \text{ and } B) = P(A) \cdot P(B|A)$$

General Addition Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Conditional Probability

$$P(B|A) = \frac{P(B \text{ and } A)}{P(A)}$$

Discrete Random Variables

Discrete Probability Distributions

Expected Value (mean)

$$\mu_x = E(x) = \sum [x_i \cdot P(x_i)]$$

Variance

$$\sigma_x^2 = \sum [x_i^2 \cdot P(x_i)] - \mu^2$$

Standard Deviation

$$\sigma_x = \sqrt{\sigma^2}$$

Binomial Distribution

$$x = \#\ of\ successes$$
$$p = probability\ of\ success$$
$$q = 1 - p$$

Binomial Probability Distribution

$$P(x) = nCx \cdot p^x \cdot q^{n-x}$$

Expected Value (mean)

$$\mu = np$$

The Normal Distribution

***Z*-score**

$$z = \frac{x - \mu}{\sigma}$$

Raw Score

$$x = z\sigma + \mu$$

***Z*-score for a Sample Mean**

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Estimation

ONE SAMPLE

Confidence Intervals for μ

σ *is known*

$$E = z_c \cdot \frac{\sigma}{\sqrt{n}}$$

σ *is unknown.*

$$E = t_c \cdot \frac{s}{\sqrt{n}}$$

$$df = n - 1$$

Confidence Intervals for p

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

$$where\ \hat{p} = \frac{x}{n}$$

TWO SAMPLES

Confidence Intervals for $\mu_1 - \mu_2$

σ_1, σ_2 *known.*

$$E = z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

σ_1 *or* σ_2 *unknown*

$$E = t_c \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

df = smaller of $n_1 - 1$ or $n_2 - 1$.

Confidence Intervals for $p_1 - p_2$

$$E = z_c \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$where\ \hat{p} = \frac{x}{n}$$

Lower Limit: point estimate – error
Upper Limit: point estimate + error

Minimum Sample Size

$$n = \left(\frac{z_c \cdot \sigma}{E}\right)^2$$

$$n = p(1-p)\left(\frac{z_c}{E}\right)^2$$

Without an estimate for p, use $p = 0.5$.

Level of Confidence, c	0.8	0.85	0.9	0.95	0.98	0.99
Critical Value, z_c	1.28	1.44	1.645	1.96	2.33	2.575

Hypothesis Testing

One Sample

Hypothesis Tests for μ

σ *is known*

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

σ *is unknown.*

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

$$df = n - 1$$

Hypothesis Tests for p

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

$$where \;\; \hat{p} = \frac{x}{n}$$

Two Samples

Hypothesis Tests for μ_1, μ_2

σ_1, σ_2 *known*

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

σ_1 *or* σ_2 *unknown*

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

df = smaller of $n_1 - 1$ or $n_2 - 1$.

Hypothesis Tests for p_1, p_2

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$where \quad \bar{p} = \frac{x_1 + x_2}{n_1 + n_2}, \quad \hat{p} = \frac{x}{n}$$

Regression

Correlation coefficient

$$r = \frac{\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)}{n - 1}$$

OR

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

Least-Squares Regression Line

$$\hat{y} = a + bx$$

where

$$b = r \cdot \frac{s_y}{s_x} = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

$$a = \bar{y} - b\bar{x}$$