

Review - Exam 1
Ch 1 - 5

Variables

1) Numeric
   a) discrete
   b) continuous

2) Categorical
   a) ranked
   b) unranked

Tables / Graphical Displays

1) Dot Plot
2) Stem and Leaf Plot
3) Box Plots
4) Bar Graph
5) Pie Chart
6) Frequency Distribution
7) Histogram

Misusing Statistics

Calculations

$\sum x$

$\sum x^2$

$(\sum x)^2$

$\sum (x - \bar{x})$

$\sum x^2 - \frac{(\sum x)^2}{n}$

$\sum (x - \bar{x})^2$

$n!$

$\binom{n}{m} = \frac{n!}{m!(n-m)!}$

b^n

ch2

R.14

Averages - Centers of Data

Mean: $\bar{x} = \frac{\sum x}{n}$

Median: $\frac{1}{2}(n+1)^{th}$
$\tilde{x}$ ranked observation

Mode: most frequent
observed value

Sample Standard Deviation

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

or

$$s = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}}$$

Interpretation

Interpretation

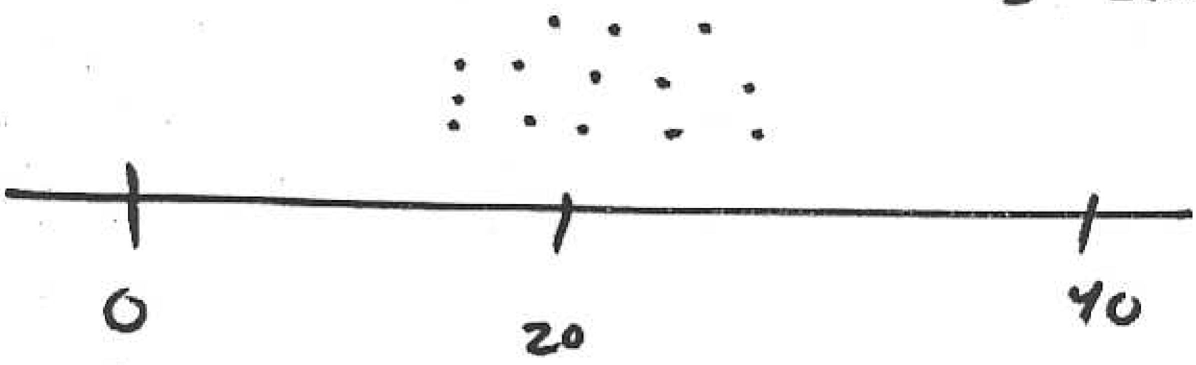
Value of s indicates how "spread out" the data is

1) $s = 0$ → No variation in the data; values all the same

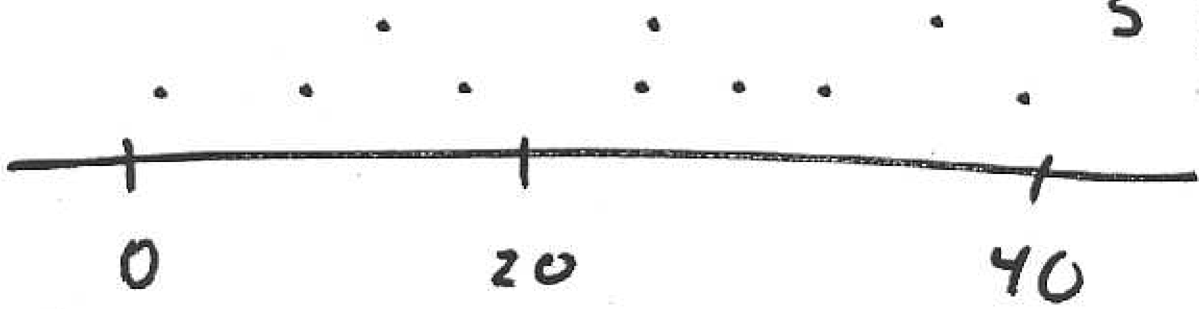
2) s "small" → the data values are not widely dispersed

3) s "large" → the data values are widely dispersed

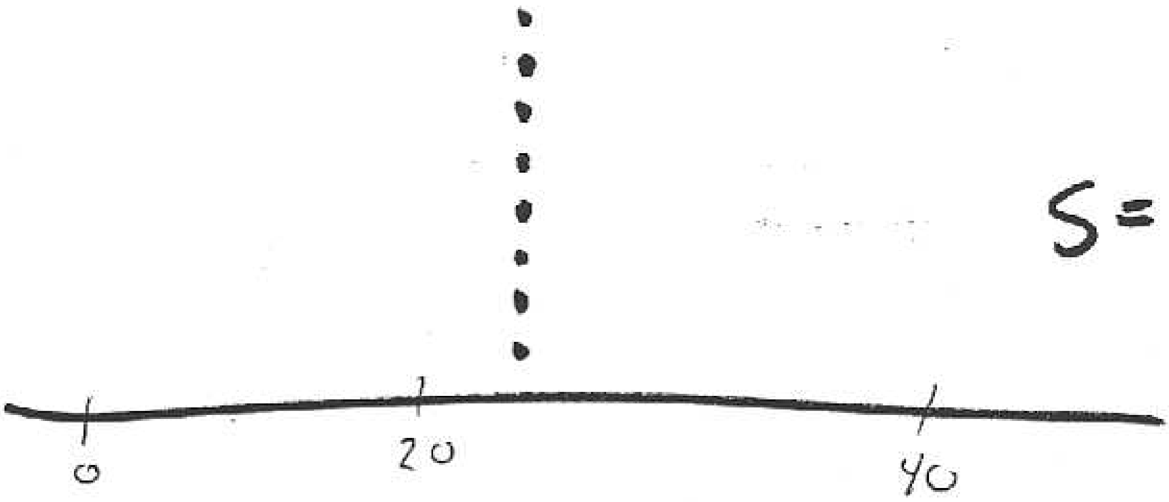
S small



S large



S=0



$$IQR = Q_3 - Q_1$$

$$\text{Skewness} = \frac{3(\bar{x} - \tilde{x})}{s}$$

Symmetric Dist.

Pos. Skewed "

Neg. Skewed "

5-Number Summary

Boxplots

ch 3 - Correlation

Pearson Corr. Coeff.

$$r = \frac{SS(xy)}{\sqrt{SS(x) \cdot SS(y)}}$$

where

$$SS(xy) = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}$$

$$SS(x) = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$$

$$SS(y) = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$$

Interpretation

1) Positive Corr.
2) Negative Corr.
3) Zero Corr.
4) $-1 \leq r \leq 1$
5) Cause and Effect

Regression Analysis

Regression (Prediction) Equation

$$Y = b_0 + b_1 x$$

$$b_1 = \frac{SS(XY)}{SS(X)}$$

$$b_0 = \bar{Y} - b_1 \bar{X}$$

Predicting Values of Y

Point Estimator: $\hat{y} = b_0 + b_1 x$

Probability - Relative Frequency Definition

Def: Suppose an experiment consists of n trials, and k of these trials result in event E. Then

$$\hat{P}(E) = \frac{k}{n} = \frac{\#\text{ successful repetitions}}{\text{total }\#\text{ repetitions}}$$

Note: This is called the empirical probability of an event or the relative frequency of the event.

Probability - Equally Likely Outcomes

Def: Suppose an experiment can result in one of m equally likely outcomes. Suppose that r of these outcomes result in event A occuring. Then the theoretical probability of event A is

$$P(A) = \frac{r}{m}$$

$$= \frac{\#\text{ outcomes in event } A}{\text{total } \#\text{ possible outcomes}}$$

Note: For each outcome in S.S.

$$P(\text{outcome}) = \frac{1}{\text{total } \#\text{ possible outcomes}}$$

A discrete probability distribution is a list (or description) of the values the random variable can have, along with the associated probabilities.

We can do this using a probability tree.

Rules

The probability of an event E is always between 0 and 1, inclusive:

$$0 \leq P(E) \leq 1$$

$P(E) = 0 \longrightarrow$ event E cannot occur

$P(E) = 1 \longrightarrow$ event E must always occur

2) The probability of event A is equal to the sum of the probabilities of the outcomes in event A

$$P(A) = \sum_{\substack{\text{all}\\ \text{outcomes}\\ \text{in } A}} P(\text{outcome})$$

Complementory Event

Def: Suppose A is an event. The complement of event A, denoted "not A", is the event "A does not occur".

Rule of Complementary Events

$$P(\text{not } A) = 1 - P(A)$$