Review Sheet

Friday, February 14, 2020 5:33 PM

1. Statistics

- samples to learn about population

observed subset, size n — complete set of items that interest investigation

Parameter

- numerical measurement that describes characteristic of population

Statistic

- numerical measure that describes a specific characteristic of a sample

Sampling Errors

- random differences between sample/population
- cancel out on average
- decrease as sample size grows

Non Sampling errors

- Systematic differences between sample/population
- Don't necessarily cancel out on average
- Don't necessarily decrease as sample size grows

Margin of Error

$$ME \approx 1.96 \cdot \sqrt{\frac{p(1-p)}{n}}$$

Usually Report $p \pm ME$

- Parameter should relate to what you are interested in

2. Types of Variables

- Nominal ($=$ or $\neq$)
- Ordinal ($<$ or $>$)

Categorical, qualitative differences

- Interval ($+$ or $-$) (No natural 0)
- Ratio ($\times$, $\div$, x^2, $\log x$ ... etc) (Absolute 0)

Numerical, quantitative difference

- discrete vs continuous

takes values from a discrete set — any value in a range

Relative Frequency

- frequency / sample size

Histogram

- plotting frequency
- can vary bin widths ← even have 2 bins w/ different widths

Measures of Central Tendency

- Mean

Sample Mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ } Sensitive to outliers

Outlier is an unusually small/large value

- Median

n is odd $med(X) = X_{\frac{n+1}{2}}$

n is even $med(X) = \frac{X_{\frac{n}{2}} + X_{\frac{n+1}{2}}}{2}$

} Mid point, less sensitive to outliers

- Quantile
  - generalization of median
  - a number $0 < \alpha < 1$ where α-quantile $= \alpha \cdot 100\%$

Percentiles, Deciles, Quintiles, Quartiles
1%, 10%, 20%, 25%

- Range
  - measure of variability / spread
  - 100% quantile − 0% quantile, max − min

} Very sensitive to outliers

- Inter Quartile Range (IQR)

$IQR \triangleq 3^{rd}$ Quartile $- 1^{st}$ Quartile

$= 75\% - 25\%$

- Variance

- Variance

$$\sigma^2 \triangleq \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

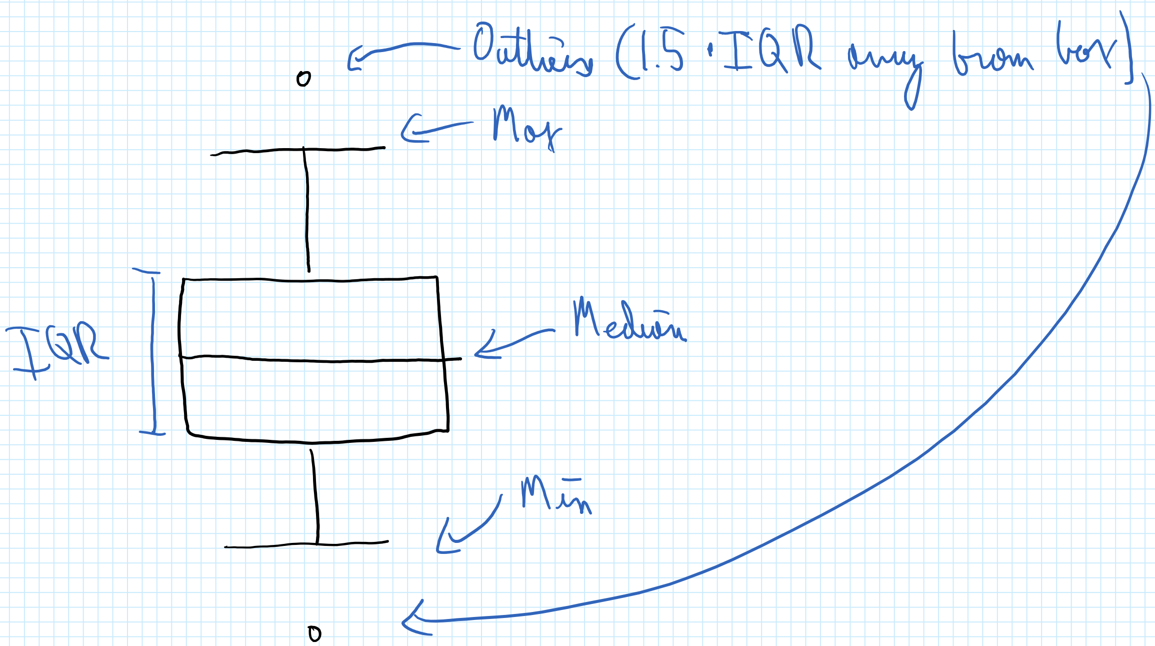
- Standard Deviation

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

- Mean Absolute Deviation

$$MAD \triangleq \frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}|$$

3. - Box Plot
   - visualize all 5 quartiles



Measures of Symmetry

$$\text{Skewness} := \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\sigma^3}$$

- Negative - Left skewed (usually mean < median)
- 0 → Symmetry
- Positive → right skewed (usually mean > median)

- Covariance

$$s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

- $-$ → negative linear dependence
- 0 → No linear dependence
- $+$ → positive linear dependence

- Correlation

$$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y}$$

$-1 \leq r \leq 1$ Same rules

- standardized and unitless

Linear Transformations

$$W_i = b \cdot x_i + a$$

$$\bar{W} = b\bar{x} + a$$

$$s_w^2 = b^2 s_x^2$$

$$s_w = |b| s_x$$

$$s_{wy} = b \cdot s_{xy}$$

$r_{wy} = \pm r_{xy}$ } sign in $\pm$ depends on sign of b

	Sample Statistic	Population Parameter
Mean	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\mu_x := E(X_i)$
Var	$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$	$\sigma^2 = E[(X_i - \mu_x)^2]$
S.D.	$s_x = \sqrt{s_x^2}$	$\sigma = \sqrt{\sigma^2}$
Cov.	$s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$	$\sigma_{xy} = E[(X_i - \mu_x)(Y_i - \mu_y)]$
Corr	$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y}$	$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$

Two Pillars of Statistics

- LLN (Law of Large #'s)
  - as sample size $\uparrow$, statistics approximate population parameter better and better
- CLT (Central Limit Theorem)

- if sample size is sufficiently large, difference between sample statistic and population parameter follows a Gaussian distribution

4. Random Experiment
   - one whose outcomes are random

Basic Outcome
   - finest grained relevant outcome of a random experiment

Sample Space (Ω)
   - set of all possible basic outcomes

Event (E)
   - subset of sample space

Occurrence
   - After experiment, only one basic outcome will happen
   - lets call it $\omega_{realized}$
   - E has "occurred" if $\omega_{realized} \in E$

Probability
   - defined on events
   - $\mathbb{P}(E)$ is the probability of E)
   - Must satisfy the 3 Axioms

1. $0 \le P(E) \le 1$
2. $P(\Omega) = 1$
3. if $E_1 \to E_N$ are mutually exclusive (empty intersect)

$$P(E_1 \cup \to E_N) = P(E_1) + \dots + P(E_N)$$

Classical Probability
   - all outcomes are equally likely

- $P(E) = \frac{\#(E)}{N}$

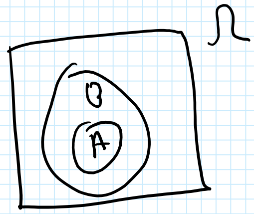
5.

Complement

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

The Inclusion Rule

- event A logically implies event B
- $A \subseteq B$



If $A \subseteq B$ then $P(A) \leq P(B)$

The Union (Logical Addition) Rule

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Conditional Probability

$$P(B/A) := \frac{P(A \cap B)}{P(A)}$$

conditional probability of B given A

- can treat $P(\cdot/A)$ as a restricted sample space
  - must follow 3 axioms

Multiplication Rule

$$P(A \cap B) = P(B/A) \cdot P(A)$$

comes from above

6. Statistical Independence

- A and B are statistically independent if

$$P(A \cap B) = P(A) \cdot P(B)$$

$$\therefore P(B) = P(B/A)$$

Pairwise Independence

$E_1, E_2 \dots E_n$ } For a list of events

$P(E_i \cap E_j) = P(E_i) \cdot P(E_j)$ } If all combos of pairs are statistically independent the list is pairwise independent

Mutual Independence

For a similar list of events as ↑

$P(E_1 \cap E_2 \cap \dots \cap E_k) = P(E_1) \cdot P(E_2) \dots P(E_k)$ } for any subset of events

Law of Total Probability

$E_1, E_2, \dots E_k$

Mutually Exclusive if $E_i \cap E_j = \emptyset$ } For all pairs in the list

Exhaustive if $E_1 \cup E_2 \dots \cup E_k = \Omega$

$\therefore$ If $E_1 \to E_k$ are mutually exclusive and exhaustive

$$P(A) = P(A|E_1) \cdot P(E_1) + P(A|E_2) P(E_2) \dots + P(A|E_k) \cdot P(E_k)$$

Bayes Rule

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

Permutations: $\frac{n!}{(n-k)!}$
(order matters)

Combinations: $\frac{n!}{k!(n-k)!}$
(order doesn't matter)

} N possibilities, K slots

7. Random Variable (X)

Ω, is a sample space

An RV, X, is a fixed function that maps each basic outcome $\omega \in \Omega$ to a real #, $X(\omega)$

can be discrete, continuous, or mixed

Realization (x)

- a particular numerical value in $\mathbb{R}$ that a RV takes on

$$\{X = x\} := \{\omega : X(\omega) = x\}$$

Support

- the set of all possible realizations of a RV

$$Supp(X) := \{X(\omega) : \omega \in \mathbb{R}\}$$

(Set of all realizations)

Probability Mass Function

Given Ω, $\mathbb{P}$, X

$$p(x) = \mathbb{P}(X = x)$$

← Get out the probability it occurs

Plug in a realization

$$0 \le p(x) \le 1$$

$$\sum_{i=1}^{n} p(x_i) = 1$$

— sum of pmf for all realizations in the support $= 1$

Cumulative Distribution (CDF)

$$F(x_0) := \mathbb{P}(X \le x_0)$$

Properties of CDF

- always increasing

$-F(-\infty)=0$

$-F(\infty)=1$

Expectation

$$\mu_X = \mathbb{E}(X) := \sum_{x \in \text{Supp}(X)} x \cdot p(x)$$

The sum of all realizations in the support of X multiplied by their probability

8. RV function on an RV?

$g(X) \rightarrow g(X(\omega))$

say $Y := g(X)$

$\text{Supp}(Y) = \{g(x_1), g(x_2) \dots g(x_n)\}$

$$p_y = P(g(X)=y) = \sum_{x : g(x)=y} p_x(x)$$

$$\mathbb{E}[Y] = \sum_{y \in \text{supp}(Y)} y\, p_y(y) \quad \text{or} \quad \mathbb{E}[Y] = \sum_{x \in \text{supp}(X)} g(x)\, p_x(x)$$

$$\boxed{E[g(X)] \neq g(E(X))}$$

only true if b is linear

$E(bx+a) = b \cdot E[x] + a$

Variance

$$\text{var}(X) = \sigma^2 = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[(X-\mathbb{E}(X))^2]$$

$$\sigma^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Standard Deviation

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

Linear Transformations?

$Var(aX+b) = a^2 Var(X)$

*Try to use definitions wherever you can

Bernoulli Trial

- RV w/ parameter p where

$p(1) = P(X=1) = p$

$p(0) = P(X=0) = 1-p$

2 outcomes, where "success" is 1 and has probability p, "failure" is 0 and has probability 1-p

$X \sim Ber(p)$

$E(Ber(p)) = p$

$Var(Ber(p)) = p(1-p)$

Consider Multiple Bernoulli Trials (n trials)

$X \sim Binomial(n, p)$

$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$ ← all of combinations $\frac{n!}{x!(n-x)!}$

probability of x "successes" in n trials

$\frac{4!}{2!(4-2)!}$ $6 \cdot \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^2$ $\frac{8}{27}$

$\mathbb{E}(X) = n \cdot p$

RV

- defined over Ω, depend on ω realized

Constants

- don't depend on ω
- $\mathbb{E}(X)$ is a constant