Acc. Precalc

Emily Thom
9-7-16

Probability & Statistics Study Guide

Unit 2

I. Combinations & Permutations

→ **Factorial** if n is a positive integer then n factorial (!) =

$$n! = n * (n-1) * (n-2) * (n-3) \ldots$$

- Zero - zero is NOT a positive integer, but

$$0! = 1$$

→ **Permutation** arrangement of things in which order is important

↳ n things taken r at a time =

$$P(n,r) = {}_nP_r = \frac{n!}{(n-r)!}$$

→ **Combination** arrangements of objects in whic order is NOT important

↳ the combination of n things taken r at a time =

$${}_nC_r = C(n,r) = \binom{n}{r} = \frac{n!}{(n-r)! * r!}$$

*combinations are smaller than permutations

II. Pascal's Triangle and Binomial Theorem

- a table of combinations
- coefficients of x and y

$(x+y)^n$

x's exponents go from n to 0
y's exponents go from 0 to n

what is the r^{th} term of $(a+b)^n$?

$$C(n, r-1)\ a^{n-r+1}\ b^{r-1}$$

III. Probability

probability: the number of ways things can happen divided by the sample space

↳ **sample space** the set of all possible outcomes

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

$A \cap B$ → intersection of A & B

→ **compliment** all outcomes in the sample space that are NOT the event

$P(A') = 1 - P(A)$

IV. Basic Statistics

- **center** mean / median
- **Central Limit Theorem (CLT)** $(n > 30)$
  1. distribution of sampling means is approximately normal
  2. mean of samping distribution = population mean
  3. standard deviation of samping means = $\frac{\sigma}{\sqrt{n}} = S_{\bar{x}}$
- **sample mean** $(\bar{x})$
  the mean of the population, a statistic measuring the average of the observations in the sample
- **sample proportion** - indicates the proportion of success in a particular sample, written as $\hat{p}$
- **population mean** (μ)
- **population standard deviation** (σ)

tells you measures of spread

- **mean absolute deviation**
  ① find the mean
  ② find difference between each value & the mean
  ③ find the absolute value of each value from #2
  ④ find the mean of the values from #3
- **variance**
  the average of the squared differences from the mean
  ① find the mean
  ② find difference between each value & the mean
  ③ square each result
  ④ find the mean of the values from #3
- **Standard deviation**
  ① find the variance
  ② find the square root of the variance
- **box plot**
  - min, Q1, mean, Q3, max
  - outliers: 1.5 (IQR) ± Q_1 / Q_3

V. Sampling

techniques
- **simple random** - completely random, every possible sample has the same chance of being selected
- **systematic** - selection in intervals / in an order
- **stratified random** - create groups & select using simple random from each group
- **cluster** - make groups & select 1 group (cluster) to survey

methods
- **convenience** - choosing those who are easiest to reach
- **voluntary response** - people choose themselves & respond to a general appeal

VI. Empirical Rule & Normal Distribution

(68-95-99.7 rule)

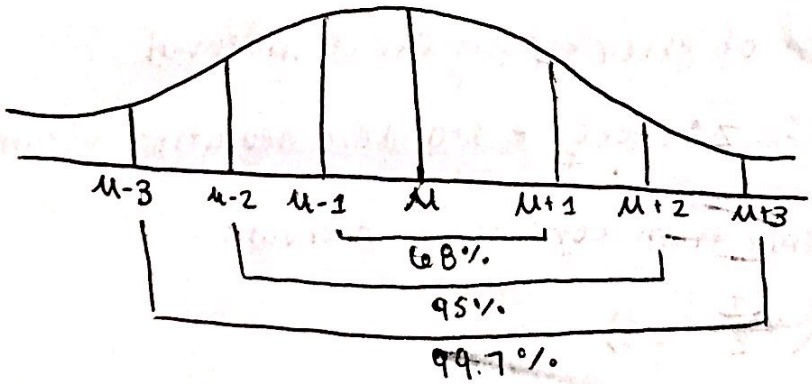
-estimates the spread of data in normal distribution given the mean & standard deviation

-normal distribution:

-symmetric, bell shaped, follows empirical rule

change in mean = shifts curve left/right

change in σ = shrinks/stretches curve up/down



→ to find probability/percentage under the curve:

normal cdf (upper value, lower value, mean (μ), standard deviation (σ))

-z scores

the number of standard deviations the x value lies above or below the mean when $\mu = 0$ and $\sigma = 1$.

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

z = zscore
x = your score
μ = mean
σ = standard deviation

VII. Binomial Distribution / Probability

conditions:

- each trial has two outcomes
- fixed number of trials (n)
- trials must be independent
- probability of success (p) must be the same for each trial

$P(x) = \binom{n}{r} (p)^x (1-p)^{n-x}$ OR $P(x) = \left(\frac{n!}{(n-x)!\,x!}\right)(p)^x (1-p)^{n-x}$

VIII. Jelly Blubber Lab

- means tend to be higher when selecting by convenience
  - ↳ eyes drawn to larger objects
- random selection is good
  - ↳ creates a smaller distribution w smaller values
- Central limit theorem
  - ↳ see IV

IX. Confidence Intervals

mean ± margin of error = confidence interval

margin of error: z* score * standard deviation of sample means ($s_{\bar{x}}$)

how to find z* score from confidence interval:

invnorm $\left(\frac{CI}{2} + .5\right)$ (Left)

- invT is used when $n < 30$

→ less confidence = smaller interval = smaller z score

→ more confidence = larger interval = larger z score

X. Proportions & Confidence Intervals

$CI = \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

where $\hat{p}$ = proportion of sample

Standard deviation of sample proportions