# MATH 460: Probability & Statistics — January_2019

## Ch 7 Tests of Hypothesis and Significance

**Class 7 Notes** — **Dr. Basilio**

Mon Feb_4 $\cup$ Tues Feb_5

* * *

### Guiding Question(s)

(1) How can we use sample data to test hypotheses (or claims) about the entire population?

## Chapter 7: Tests of Hypotheses and Significance

## Hypothesis Testing

Previously we used confidence intervals to estimate unknown population parameters. Now we are going to look at a more formal method for testing whether a given value is a reasonable value of a population parameter. To do this we need to have a hypothesized value of the population parameter.

The general idea of hypothesis testing involves:

1. Making an initial assumption.
2. Collecting evidence (data).
3. Based on the available evidence (data), deciding whether to reject or not reject the initial assumption.

Every hypothesis test–regardless of the population parameter involved–requires the above three steps.

We will compare data from a sample to a hypothesized parameter. In each case, we will compute the probability that a population with the specified parameter would produce a sample statistic as extreme or more extreme to the one we observed in our sample. This probability is known as the **P-value** and it is used to evaluate statistical significance.

### Example 1: Hypothesis Testing

(a) Claim: "Most people get their jobs from networking"

We let $p$ be the proportion of people from a population of working adults who got their jobs from networking. We want to check the claim about the entire population. What is "most"? We should say that over 50% should mean most. So the hypothesis we want to check is translated to: $p > 0.50$.

(b) Claim: "The average monthly rent of a one-bedroom apartment in Anaheim, CA is less than $1,100$?

We clearly see that the population mean is what we are making a statement about. We can state this hypothesis as: $\mu < 1,100$.

## Example 2: Criminal Trial Analogy

One place where you can consistently see the general idea of hypothesis testing in action is in criminal trials held in the United States. Our criminal justice system assumes "the defendant is innocent until proven guilty." That is, our initial assumption is that the defendant is innocent.

In the practice of statistics, we make our initial assumption when we state our two competing hypotheses–the null hypothesis ($H_0$) and the alternative hypothesis ($H_A$). Here, our hypotheses are:

- $H_0$: Defendant is not guilty (innocent)
- $H_A$: Defendant is guilty

In statistics, we always **assume the null hypothesis is true**. That is, the null hypothesis is always our initial assumption.

The prosecution team then collects evidence–such as finger prints, blood spots, hair samples, carpet fibers, shoe prints, ransom notes, and handwriting samples–with the hopes of finding "sufficient evidence" to make the assumption of innocence refutable.

In statistics, the **data** are the evidence.
The jury then makes a decision based on the available evidence:

- If the jury finds sufficient evidence–beyond a reasonable doubt–to make the assumption of innocence refutable, the jury **rejects the null hypothesis** and deems the defendant guilty. We behave as if the defendant is guilty.
- If there is insufficient evidence, then the jury **does not reject the null hypothesis**. We behave as if the defendant is innocent.

In statistics, we always make one of two decisions. We either "reject the null hypothesis" or we "fail to reject the null hypothesis." However, we never *prove* that the null hypothesis is true. In other words, there's always room for error.

## Definition 1: Null and Alternative Hypotheses

- **HYPOTHESIS TESTING:** testing whether or not a claim or hypothesis is valid.
- **NULL HYPOTHESIS:** is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value. The null hypothesis is denoted by $H_0$.
- **ALTERNATIVE HYPOTHESIS:** is a statement that the population parameter has a value that somehow differs from the null hypothesis. The alternative hypothesis is denoted by $H_A$ (or $H_1$).

  For our purposes, the symbolic form of the alternative hypothesis must use one of these symbols: $<$, $>$ or $\neq$.
- **SIGNIFICANCE LEVEL:** is the probability value used as the cutoff for determining when the sample evidence constitutes significant evidence *against the null hypothesis*. By its nature, the significance level, $\alpha$, is the probability of mistakenly rejecting the null hypothesis when it is true:

$$\text{significance level } \alpha = P(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$$

- **CRITICAL REGION (or REJECTION REGION):** is the area corresponding to all values of the test statistic that cause us to reject the null hypothesis.

## Example 3: Null and Alternative Hypotheses

Continued from Example 1.

(a) The associated hypotheses for Example 1 part (a) are:

$$H_0 : p = 0.5, \qquad H_A : p < 0.5, \quad H_A : p > 0.5, \quad H_A : p \neq 0.5$$

Notice there's one Null and three Alternative hypotheses.

(b) The associated hypotheses for Example 1 part (b) are:

$$H_0 : \mu = 1100, \qquad H_A : \mu < 1100, \quad H_A : \mu > 1100, \quad H_A : \mu \neq 1100$$

Notice there's one Null and three Alternative hypotheses.

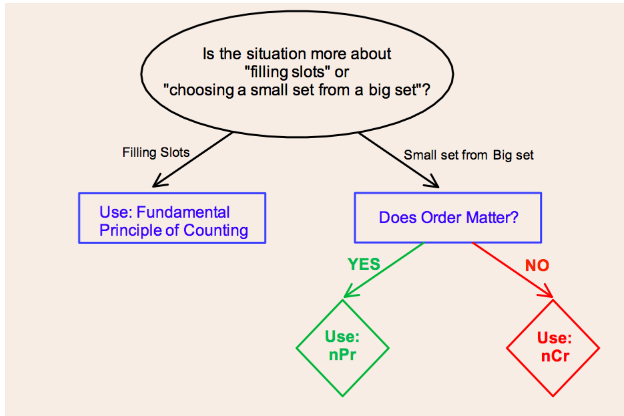
# FINAL EXAM REVIEW TOPICS

- SET THEORY
  - sets, elements, universe/sample space, subsets, disjoint, emptyset
  - union, intersections, complements, difference
  - Venn Diagrams
- PROBABILITY
  - random experiments
  - sample space, events
  - Classical approach to probability: $\boxed{P(E) = \frac{\text{\# outcomes in } E}{\text{\# outcomes in } S}}$
  - Experimental approach to probability: $\boxed{P(E) = \frac{\text{\# observations of } E}{\text{\# of total repetitions of experiment}}}$
  - AXIOMS of probability:
    - **Axiom 1** For every event $A$, then $P(A) \geq 0$
    - **Axiom 2** For the sample space $S$, then $P(S) = 1$
    - **Axiom 3** If two events $A$ and $B$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
  - Theorems of probability:
    - $P(A) + P(A') = 1$
    - $P(A') = 1 - P(A)$
    - $P(\emptyset) = 0$
    - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
  - Conditional probability: $P(A \cap B) = P(A) * B(B|A)$
- COUNTING TECHNIQUES
  - Fundamental Principle of Counting: (e.g. "picking outfits")

$$\underset{\text{\# of choices of item 1}}{\underline{\qquad\qquad\qquad}} \times \underset{\text{\# of choices of item 2}}{\underline{\qquad\qquad\qquad}} \times \underset{\text{\# of choices item 3}}{\underline{\qquad\qquad\qquad}} \times \cdots = \underset{\text{TOTAL \# of ways to combine}}{\underline{\qquad\qquad\qquad}}$$

  - Permutations: ${}_nP_r = \#$ ways of arranging $r$ objects from a collection of $n$ in order without replacement
    Keys: order matters and without replacement (e.g. making words from letters, putting books on self, etc)

- Combinations: ${}_nC_r = \#$ ways of arranging $r$ objects from a collection of $n$ without caring about order and without replacement
  Keys: order doesn't matter and without replacement (e.g. picking books to donate, selecting people to do a task (all treated equally), etc)
- How to choose method:




- TYPES OF DATA
  - Frequency Distributions, Relative Frequency Distributions, Tallies
  - Bar graphs vs Histograms
  - Measures of Center: Mean, Median, Mode, Midrange
  - Measure of Dispersion: standard deviation
  - Five Number Summary: min, Q1, median, Q3, max
  - Box-Whisker Plot
  - Skewness: positively skewed, negatively skewed, symmetric
- RANDOM VARIABLES
  - Random Variable, Discrete vs Continuous
  - be able to write the definition of a random variable (e.g. let $X = \#$ of Heads in a coin flip)
  - probability distribution of a random variable
- EXPECTATION
  - mathematical expectation
  - expectation for uniform distributions
  - Fair Game
- BINOMIAL DISTRIBUTION
  - Conditions for binomial: discrete random variable $X$ must have:
    1) fixed $n$ trials, 2) trials are independent, 3) each trial must have exactly two outcomes ('success' or 'failure'), 4) probability of success remains the same in each trial.

    $$P(X = x) = \text{ probability of } \textbf{exactly } x \text{ successful trials out of } n = \textit{binompdf}\,(n, p, x)$$

    $x = 0, 1, 2, 3, \ldots, n$
  - Examples of Binomial distribution: flipping coins (H or T), having a kid (M or F), games played in soccer (W or L), etc
  - Be able to compute: $P(X = x)$, $P(X \leq x)$, $P(X \geq x)$, $P(a \leq X \leq b)$ using either *binompdf* or *binomcdf*
- PROBABILITY DENSITY FUNCTIONS
  - probability density function (pdf) associated to a random variable
  - Probability = area under the pdf
  - Probability of exact value is zero! Thus, Must compute range of values $P(a \leq X \leq b)$

- NORMAL DISTRIBUTION
  - a continuous random variable $X$ is normally distributed if it's pdf is a bell-shaped curve
  - Normal distribution: lots of them! Bell shaped curves
    Depend on mean ($\mu$) and standard deviation ($\sigma$)
  - Properties: symmetric, continuous on $\mathbb{R}$, mean=median=mode in the center,
    total area under curve =1, area of half the normal curve = 0.5
  - USUAL GOAL: if we assume RV is normally distributed, can compute probabilities $P(a \leq X \leq b)$ using normalcdf(a,b,$\mu$, $\sigma$).
  - EMPIRICAL RULE: 68-95-99.7
    If $X$ is a normally distributed random variable, then
    * $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68$
    * $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95$
    * $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.997$

    Key: you can use empirical rule to compute probabilities ONLY IF your range of values hits exactly $\pm\sigma$, $\pm 2\sigma$, $\pm 3\sigma$, or if it includes $0.5$
  - STANDARD NORMAL: normal curve with $\mu = 0, \sigma = 1$
  - INVERSE NORMAL: reverses the process! We are given a probability $p$ and we want to find $x$ such that $P(X \leq x) = p$. Use: $x =$invNorm(p,$\mu$, $\sigma$).
- POISSON DISTRIBUTION
  - conditions for poisson: discrete random variable $X$ must have:
    (1) (possibly infinite) non-negative values, i.e. values $x = 0, 1, 2, \ldots$, (2) $\#$ successes in disjoint time intervals are independent, (3) the probability of success during a small time interval is proportional to the entire length of the time interval, (4) $\lambda$ is the mean of the number of successes in the given time interval,

    $$P(X = x) = \text{ probability of } \textbf{exactly } x \text{ successful trials out of } n = \mathit{poissonpdf}(\lambda, x)$$
  - use for problems for time intervals!
  - If $n$ is an interval of time, and $p$ is the probability of success, then $\lambda = np$.
  - A Poisson Distribution is often used to approximate the Binomial Distribution, when $n$ is "large" and $p$ is "small" (general rule is $n \geq 30$ and $p < 0.05$)
- CONFIDENCE INTERVALS
  - inferential statistics
  - population vs sample
  - parameter vs statistic
  - assumption: $n > 30$
  - confidence interval
  - Critical Level $C = 1 - \alpha$
  - Critical value $z_{\alpha/2}$: $z_{\alpha/2} =$ invNorm$((1 + C)/2)$
  - CONFIDENCE INTERVAL FOR PROPORTION $p$:
    The population proportion $p$ is within the following interval with a confidence level of $C$:
    $\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$
  - CONFIDENCE INTERVAL FOR MEAN $\mu$:
    The population mean $\mu$ is within the following interval with a confidence level of $C$: $\left(\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$