Biostatistics Review
CPH Examination

CDR LaRee Tracy, MA, PhD
Mathematical Statistician, US FDA
Laree.Tracy@fda.hhs.gov

Biostatistical Principles on CPH Exam

1. **Visualizing Data**
   A. Data Presentation
   B. Kaplan Meier
   C. Simple Regression Lines
2. **Descriptive Statistics**
   A. Central Tendencies and Variability
   B. Frequency
   C. Percentiles and Standardized Scores
3. **Statistical Probability Distributions**
   A. Normal
   B. T Distribution
   C. Binomial
   D. Chi-square
   E. Poisson
   F. F

Biostatistical Principles on CPH Exam

4. **Types of Variables and Measurement Scales**
   A. Qualitative versus Quantitative Variables
   B. Confounding
   C. Effect Modifiers
   D. Independent versus Dependent Variables
   E. Measurement Scales
5. **Measurement:** Reliability and Validity
6. **Estimation**
   A. Sampling Theory and Central Limit Theorem
   B. Estimation of Population Parameters
7. **Testing Hypotheses**

Biostatistical Principles on CPH Exam

8. **Concepts of Probability**
   A. Statistical Test Assumptions
   B. Level of Significance
   C. Decision Errors and Statistical Power
   D. Tests for group means (e.g., Z-test, t-test, ANOVA)
   E. Tests for Proportions (e.g., Chi-Square, Tests of Independence)
   F. Goodness of Fit Tests
9. **Risks and Rates:** Odds Ratio, Relative Risk
10. **Correlation and Prediction Methods**
   A. Correlation
   B. Simple Linear and C. Multiple Regression
   D. Logistic Regression
   E. Survival Analysis
11. **Interpreting Statistical Test Results**
12. **Confidence Internals**

Definitions

- **Population: largest collection of entities for which we have an interest at a particular time**

- **Population parameters: Unknown parameters that represent the theoretical value for the measure in the entire population**

- **Sample: A representative subgroup of the population**

- **Statistic: An observed characteristic of a <u>sample</u> used to estimate population parameters**

In statistics……. we are interested in drawing conclusions (inferences) about the population of interest based on findings from the sample

Data Classification and Descriptive Methods

Classification Types

- **Qualitative** ➔ non-numeric; values driven by data type
  - Examples: Gender, Race, Tumor stages
  - Recorded as a set of codes: 0/1, 1-4, I-III

- **Quantitative** ➔ numeric; differences between values have meaning on a numerical scale
  - Examples: Age, HAM-D score

Qualitative Variables

- **Categorical (or nominative)**
  - Arbitrary numbers representing names for each outcomes (level), e.g. Cause of death: 1, 2, 3, 4
  - Binary variables: used to represent two levels, e.g. alive/death: 0/1

- **Ordinal**
  - Have a natural order for levels (differs from categorical)
  - Ex: SES, level of severity

Quantitative Variables

- **Discrete** - Variables taking integer values
  - Examples: Number of Pregnancies, CAD scores, Likert Scale

- **Continuous** - Variables falling on a continuum of values
  - Examples: Blood pressure, % change from baseline in depression severity score as measured via a standardized assessment tool

How to describe data?

1. Descriptive Statistics (organizing, tabulating, graphing and summarizing data)

2. Inferential Statistics (drawing conclusions on population using data from sample)

Visualizing Data in Figures

- Dichotomous and categorical data
  - Frequencies and relative frequency charts
  - Bar charts
- Ordinal data
  - Frequencies, relative frequencies, cumulative frequencies, histograms
- Continuous data
  - Histogram, boxplots, stem and leaf
  - Measures of central tendency

Tabulating Dichotomous Data

Patient	Gender	Gender-Code
1	Male	0
2	Female	1
3	Female	1
4	Male	0

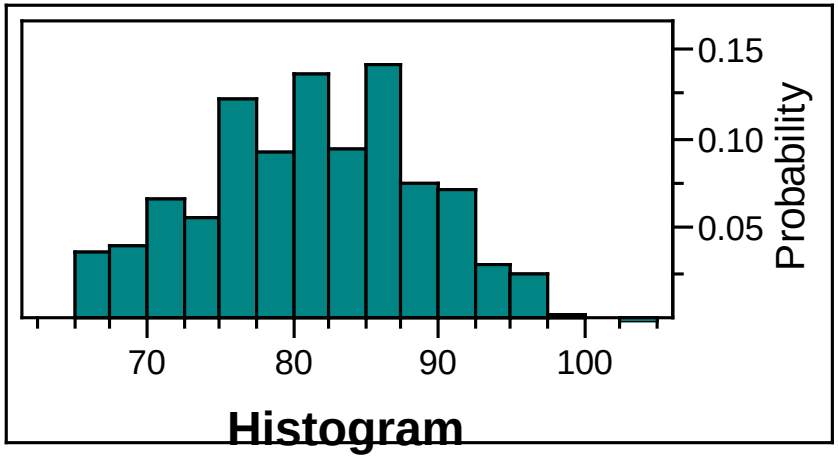
Tabulating Count Data

Patient	# of Lesions
1	12
2	2
3	3
4	8
5	5

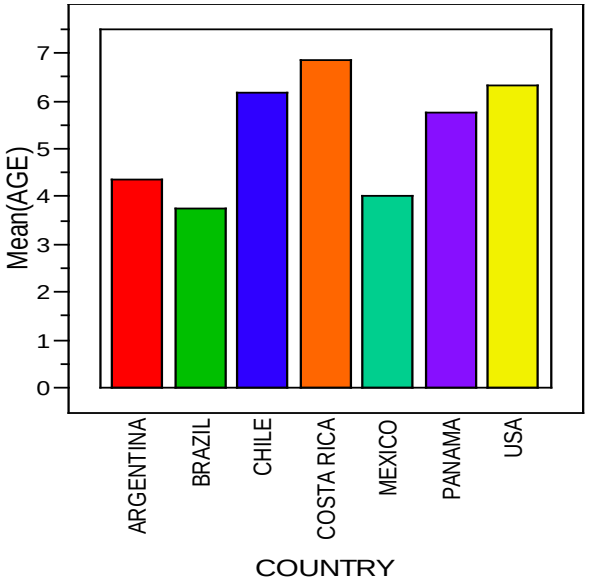
Tabulating Continuous Data

Patient	Age (yrs)
1	35.9
2	21.3
3	12.3
4	42.3
5	55.6

Graphical Data Displays



Bar chart

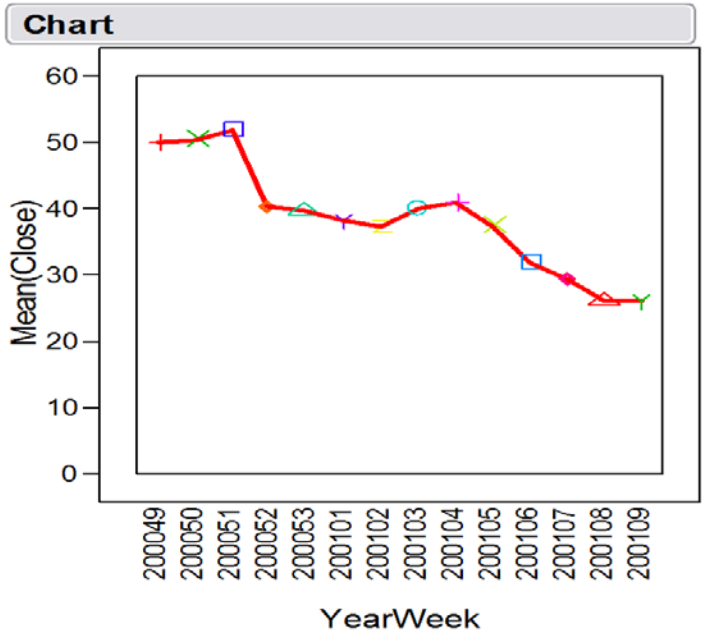


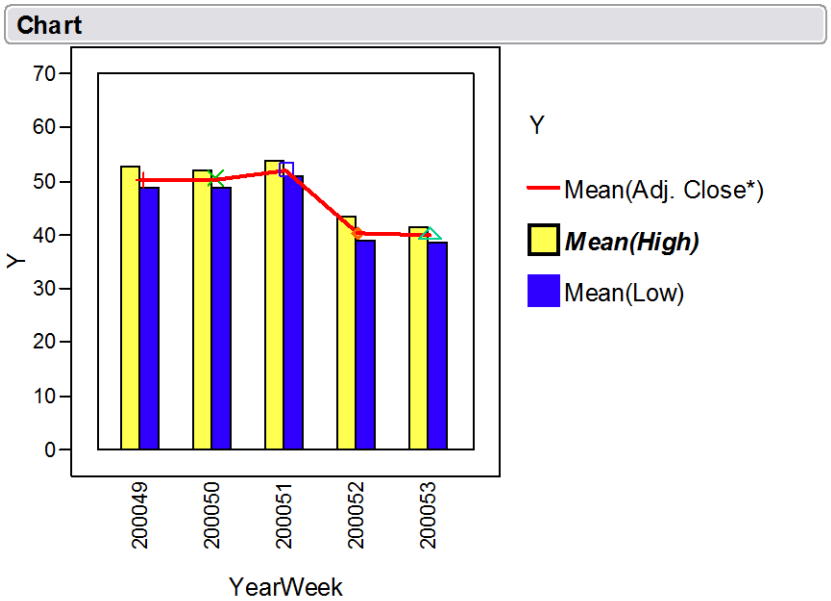
Stem-and-Leaf Display

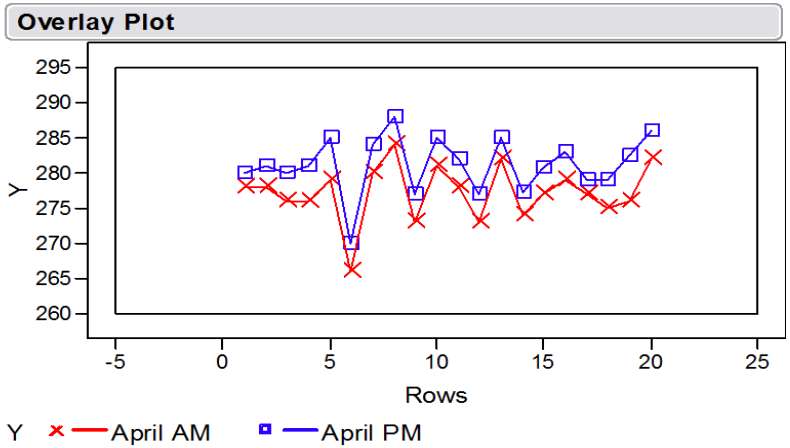
Stem	Leaf	Count
10	4	1
9	5555555566666667788	19
9	0000000000000000000000000111111111111122222222223333333333344444444444	70
8	5555555555555555555555555555555555555566666666666666666666666666666666677777777777777777777777777888888888888888888888888888888999999999999999999999999	147
8	0000000000000000000000011111111111111111111111111222222222222222222222222222222222222222222333333333333333333333333333333333333444444444444444444444444444444444	156
7	5555555555555555555555555566666666666666666666666667777777777777777777777777777777777888888888888888888888888888888999999999999999999999999999999999999	146
7	0000000000000000001111111111122222222222222222333333333333333333344444444444444444444	83
6	5555555666666666667777777788888888888888888999999999999	52

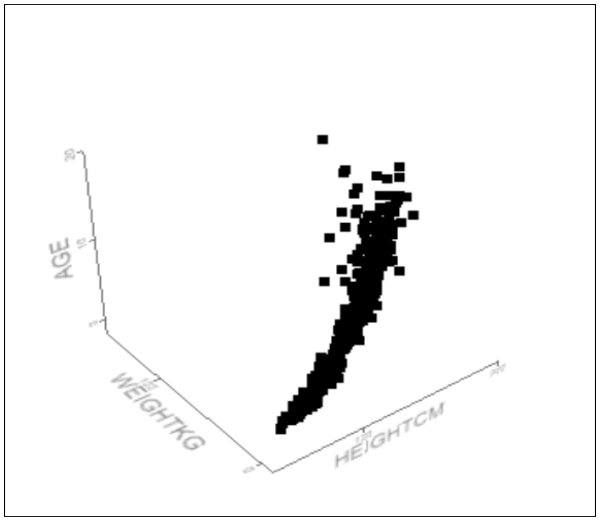
6|5 represents 65

Graphical Data Displays

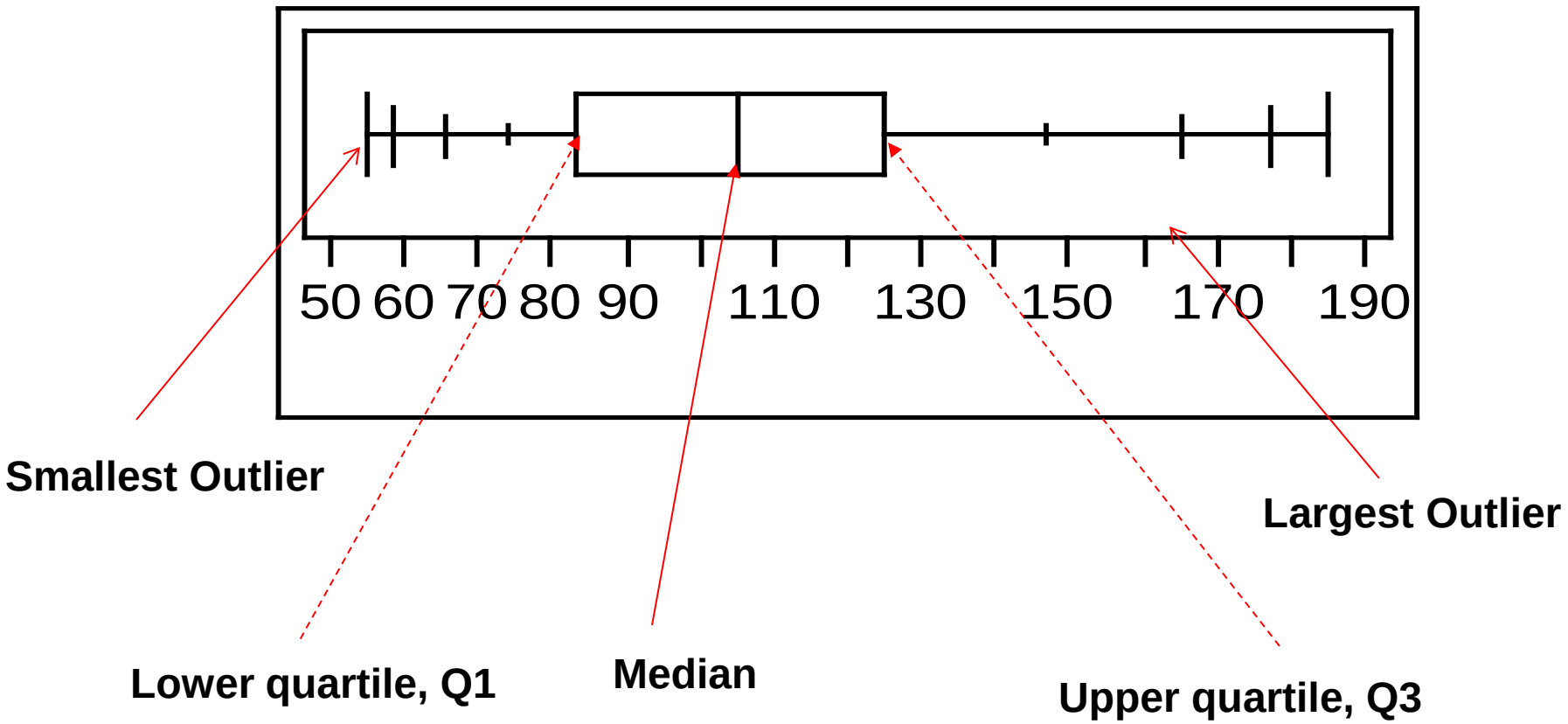








Box and Whisker Plot (boxplot)



Useful tool to identify the degree of symmetry in the data

Summarizing Data
Measures of Central Tendency

- Also called measures of the middle or of the "center"

- **Three most common:**
  - Mean: Arithmetic average of observations
  - Median: The middle observation
  - Mode: The most frequent observation

Sample Mean

- Simple arithmetic average or sum of the values divided by the number of items

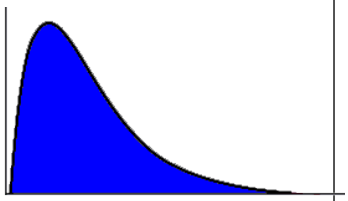
$$\overline{X} = \frac{\sum X_i}{n}$$

Example: X_i=10, 12, 11, 15, 19

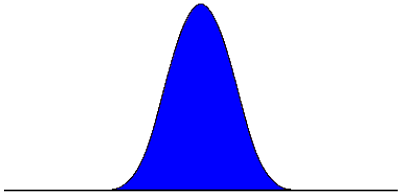
Sample mean=(10+12+11+15+19)/5=13.4

Median

- Divides the bottom 50% of the data from the top 50%
- Appropriate measure for skewed data



- Approximately equal to mean when distribution is symmetric



Summarizing Data
Measures of Variation

- Also called '***measures of spread***'
- Most commonly used:

  - Sample range
  - Sample variance
  - Sample standard deviation
  - Sample interquartile range
  - Sample coefficient of variation

Sample Range

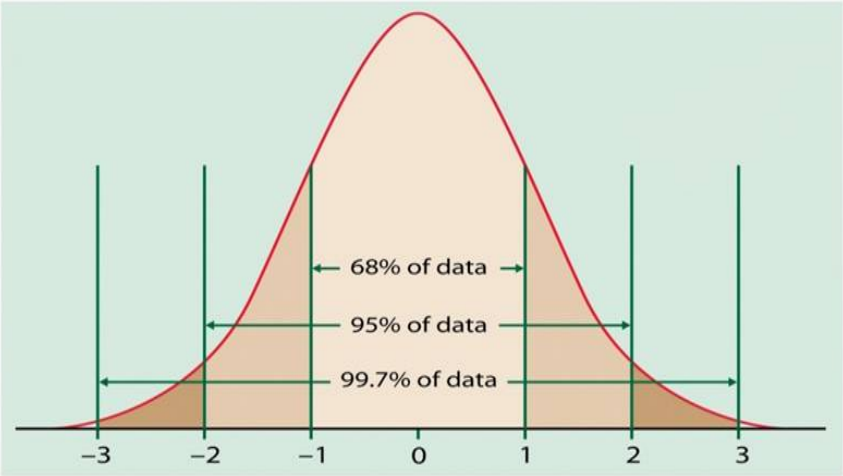
- Difference between maximum and minimum observed values
  - R=largest obs (X_{max})-smallest obs (X_{min})
  - Provides a measure of the data's variation or dispersion

- Distorted by outlier observations
  - EX: data point: 12.1, 15.4, 33.4, 65.3, 9.0
  - Range: 65.3-9.0 = 56.3

- Increases as sample size (n) increases
  - A 'rare' extreme value more likely as n increases

Sample Standard Deviation (s)

- Square root of variance ($S^2 = \frac{\sum(x_i - \overline{X})^2}{n-1}$)
- Only applicable for normally distributed data
- Gives sense of 'spread' of values around mean
- Range of ±1 SD above/below mean includes 68.2% of the data
- ±2 SD above/below mean includes 95.4% of the data
- ±1 SD above/below mean includes 99.7% of the data



Ex: Sample Variance and SD

X_i	$\overline{X}$ (Sample Mean)	$X_i - \overline{X}$	$\left(X_i - \overline{X}\right)^2$
14	10	4	16
12	10	2	4
12	10	2	4
10	10	0	0
8	10	-2	4
6	10	-4	16
Total		2	44

$$s^2 = 44/(6-1) = 8.8$$

$$s = \sqrt{\sigma^2} = \sqrt{4.9} = 2.1$$

Note that s^2 and s are estimates of population variance (σ^2) and SD (σ)

Standard Deviation- Example 2

- Trial of patients w/MDD. The mean (SD) change from baseline in HAM-D score of 0.34 (0.06)
  - 1 SD below mean is 0.34-0.06=0.28
  - 1 SD above mean is 0.34+0.06=0.41
  - +/-1 SD will include 68.2% of values, so 68.2% of values will fall between 0.28 and 0.41
  - Similarly, 95.4% of values will fall between 0.22 and 0.47 (this is the 95% CL)

Coefficient of Variation (COV)

- 100% x (SD/mean)
- A standardized measure of variation accounting for mean
- Used to compare dispersion for different variables (or samples) with different units of measure

- Questions:
  - As sample size increases, does SD increase or decrease? Decreases
  - How does sample size influence the COV? Inversely
  - How does sample size influence the confidence interval? Larger SS leads to narrower CI

Interquartile Range (IQR)

- Preferred measure of variation when <u>median</u> is used as measure of center
  - Not affected by outliers or sample size
- Difference between the first and third quartiles
  - IQR = Q3 - Q1
- **Q1** = 25th percentile; divides bottom 25% of data from top 75%
- **Q3** = 75th percentile; divides bottom 75% of data from the top 25%

Measurement

- Validity. How well does the test measure what it is supposed to measure? Does antibody positivity mean that the individual will transmit infection?

- Reliability. If you repeat the test on the same person, will you get the same result?

Screening

- Sensitivity -the probability that the test result will be positive when administered to persons who truly have the antibody (i.e., true positives / all with antibody). Algebraically, sensitivity = A/(A+C).

- Specificity -the probability that the test result will be negative when administered to persons who truly do not have the antibody (i.e., true negatives / all without antibody). Algebraically, specificity=D/(B+D).

- Predictive-value positive (PVP) - the probability that a person with a positive screening test result truly has the antibody (i.e., true positives / all with positive test). Algebraically, PVP = A/(A+B).

- Predictive-value negative (PVN) - the probability that a person with a negative screening test result truly does not have the antibody (i.e., true negatives /all with negative test). Algebraically, PVN = D/(C+D).

NOTES ON SENSITIVITY AND SPECIFICITY

Test result	Actual antibody status: Present	Actual antibody status: Absent	Total
Positive	True positive (A)	False positive (B)	All positive tests (A+B)
Negative	False negative (C)	True negative (D)	All negative tests (C+D)
Total	All with antibody (A+C)	All without antibody (B+D)	Total (A+B+C+D)

Computation of Sensitivity & Specificity

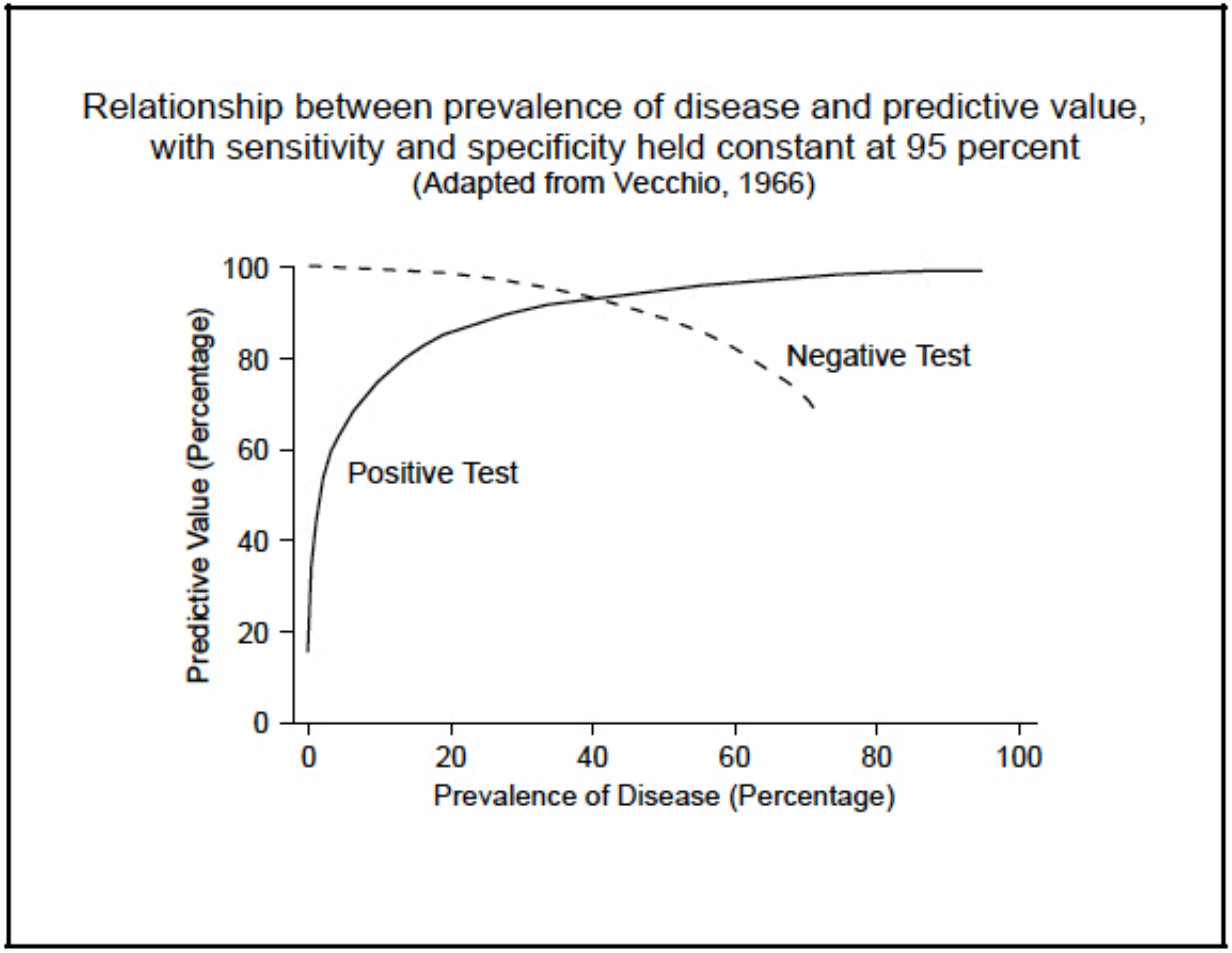
	True Diagnosis		
Test Result	Diseased	Not Diseased	Total
Positive	350	1900	2250
Negative	150	7600	7750
Total	500	9500	10,000

Sensitivity: a/a+c = 350/500=70%

Specificity: d/b+d=7600/9500=80%

Disease Prevalence: 500/10,000= 5%

Screening-Impact of Prevalence



Basics of Probability

Probability

- If a hypothetically repeatable experiment were performed an infinite number of times, the probability is the proportion of times that a particular outcome would occur

- Classic example: rolling a dice an infinite number of times, one would expect that each number would occur 1/6 of the time

Probability

- **Trial** - Collection of <u>one</u> data point

  flip a coin once, observe a head or tail

- **Experiment** - Series of trials

  observe # of heads in 10 flips of a coin

- **Event** - Single outcome or set of outcomes - What we're interested in

  getting 3 heads in 10 flips of a coin

  capital letters represent events: A, B . . .

Frequency Definition of Probability

(Suppose an experiment can be repeated many times and that one or more outcomes can result from each trial)

Probability of Outcome A:

$$P(A) = \frac{\text{\# of times A occurred}}{\text{\# of trials conducted}}$$

Probability Distributions

- Discrete Probability Distributions
  - Binomial Distribution
  - Poisson Distribution
- Continuous Probability Distributions
  - Normal Distribution
  - Standard Normal Distribution

Binomial Distribution

- Possible outcomes for each trial are binary (eg. yes or no, positive or negative)

- Useful for answering questions about the probability of X number of occurrences in n independent trials when there is a constant probability p of success on <u>each</u> trial.

Poisson Distribution

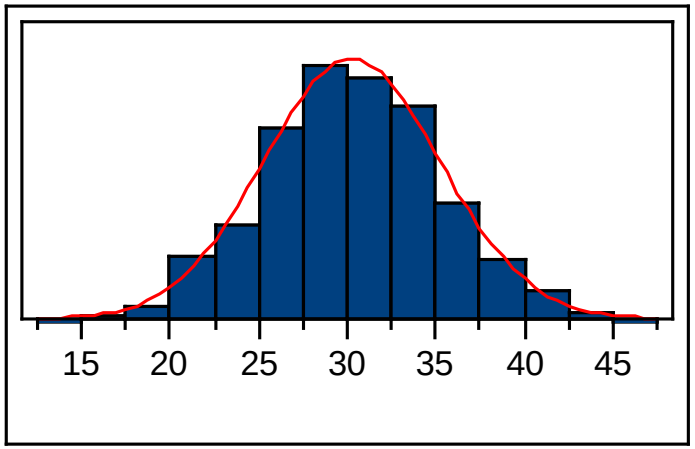
- Used for counts of events in an interval of space or time, e.g. number of infections/subject, # deaths/month

- X is the number of events in time/space interval, λ the average number of occurrences of the event in the interval

$$f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Normal Distribution

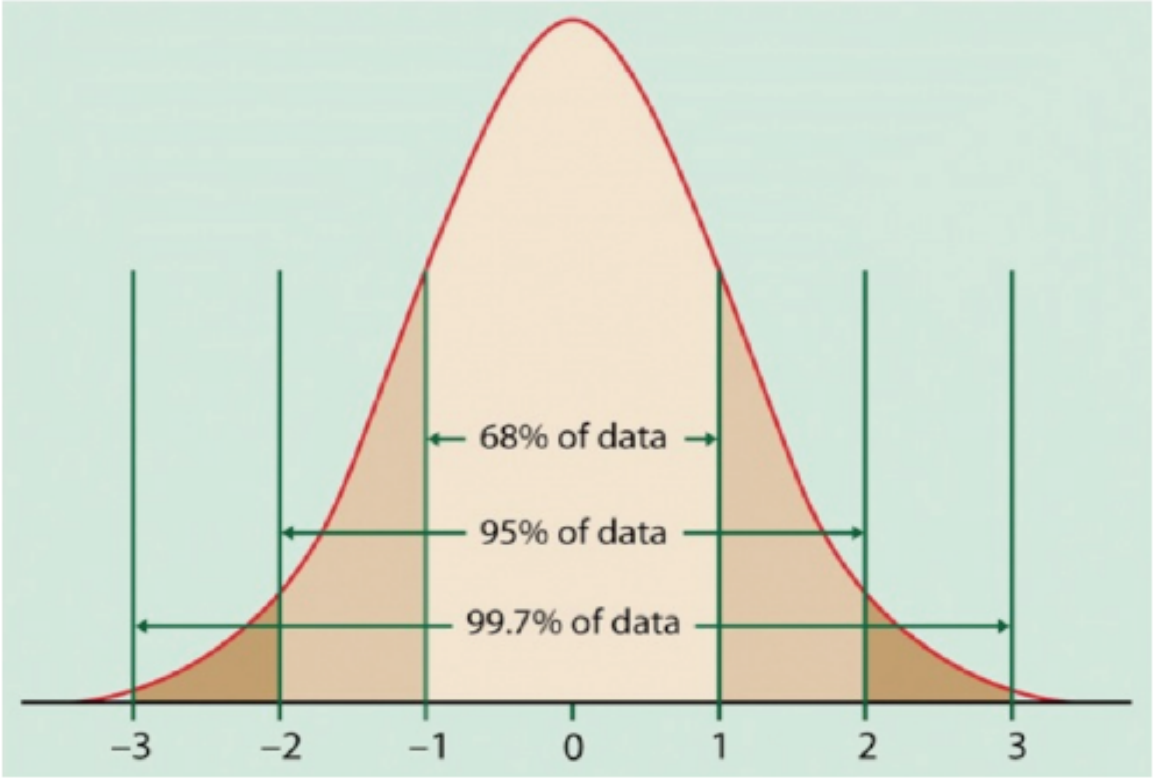
- Most commonly used distribution
- 2 parameters, mean (μ) and SD (σ)
- Continuous distribution - can take on any value, not just integers
- Bell shaped curve, symmetric about the mean



Mean=30, SD=5

Standard Normal Distribution

- Z = Standard normal (mean=0, SD=1)
- If $X \sim$Normal(μ, σ), then $Z = (X - \mu)/\sigma$ has a standard normal distribution, i.e. $Z \sim$Normal(0,1)
- Z-probabilities are available in tables (Appendix Table A2)

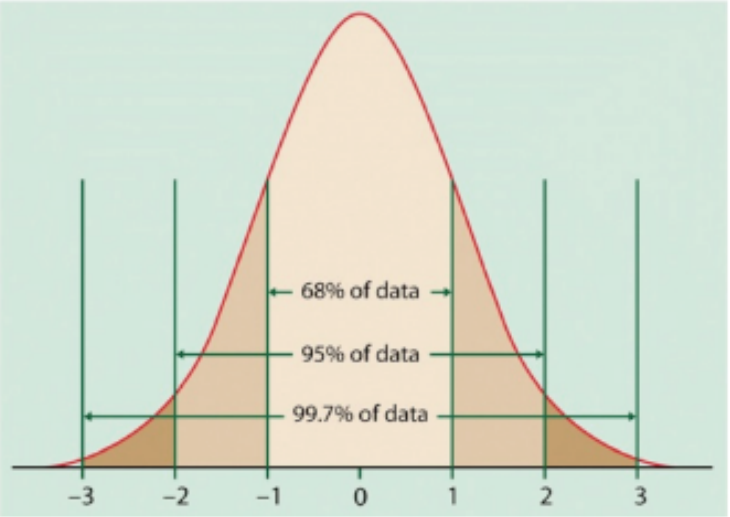


Example:

- Assume systolic blood pressure is distributed according to a normal distribution, with mean 117 and SD=17 mm Hg.
- What is proportion of population with SBP > 140mm?

$$\Pr(X > 140) = \Pr\left(\frac{X - 117}{17} > \frac{140 - 117}{17}\right)$$
$$= \Pr(Z > 1.35) = 0.089$$

- From Table A2, Pr(z>1.35) = 0.089 (one-tailed probability)
- Only about 8.9% of the population has SBP > 140 mm.



Inferential Statistics

- *A priori* stated plan to compare/assess main outcome between or among groups in a study or trial
- Objective is to draw conclusions (inferences) on population from which the sample was drawn
  - *Ex: Draw a random sample of patients with PTSD from the overall population of individuals meeting criteria for PTSD and study effects of new treatment*
- Requires certain statistical criteria are met based on sample size, amount of uncertainty you are willing to accept (type I error) and level of statistical significance.

Statistical Hypothesis

- An *a priori* set of conditions for which a study or trial is designed to prove/disprove
- **Null hypothesis (H0):** the condition which you hope to disprove by performing the trial or study
  - EX: Remission rate in new drug arm = remission rate in placebo arm (Note: This is a two-sided condition)
- **Alternative Hypothesis (H1):** opposite condition from H0
  - EX: Remission rate in new drug ≠ remission rate in placebo arm

- Statistical testing performed to either <u>reject</u> or <u>fail to reject</u> the null (note: Failing to reject the null IS NOT THE SAME as accepting it. Means have too little evidence to refute it.)

Statistical Terms: α and β

- Type I error (α) - the probability of falsely detecting a difference when none actually exists
  - Falsely rejecting a true null hypothesis
  - Generally set at 2-sided, 0.05 (5%)

- Type II error (β) - probability of being unable to detect a true difference
  - Failing to reject a false null hypothesis
  - Generally set at 0.20 or less

- Why is α less than β? What are situations where the reverse might be true?

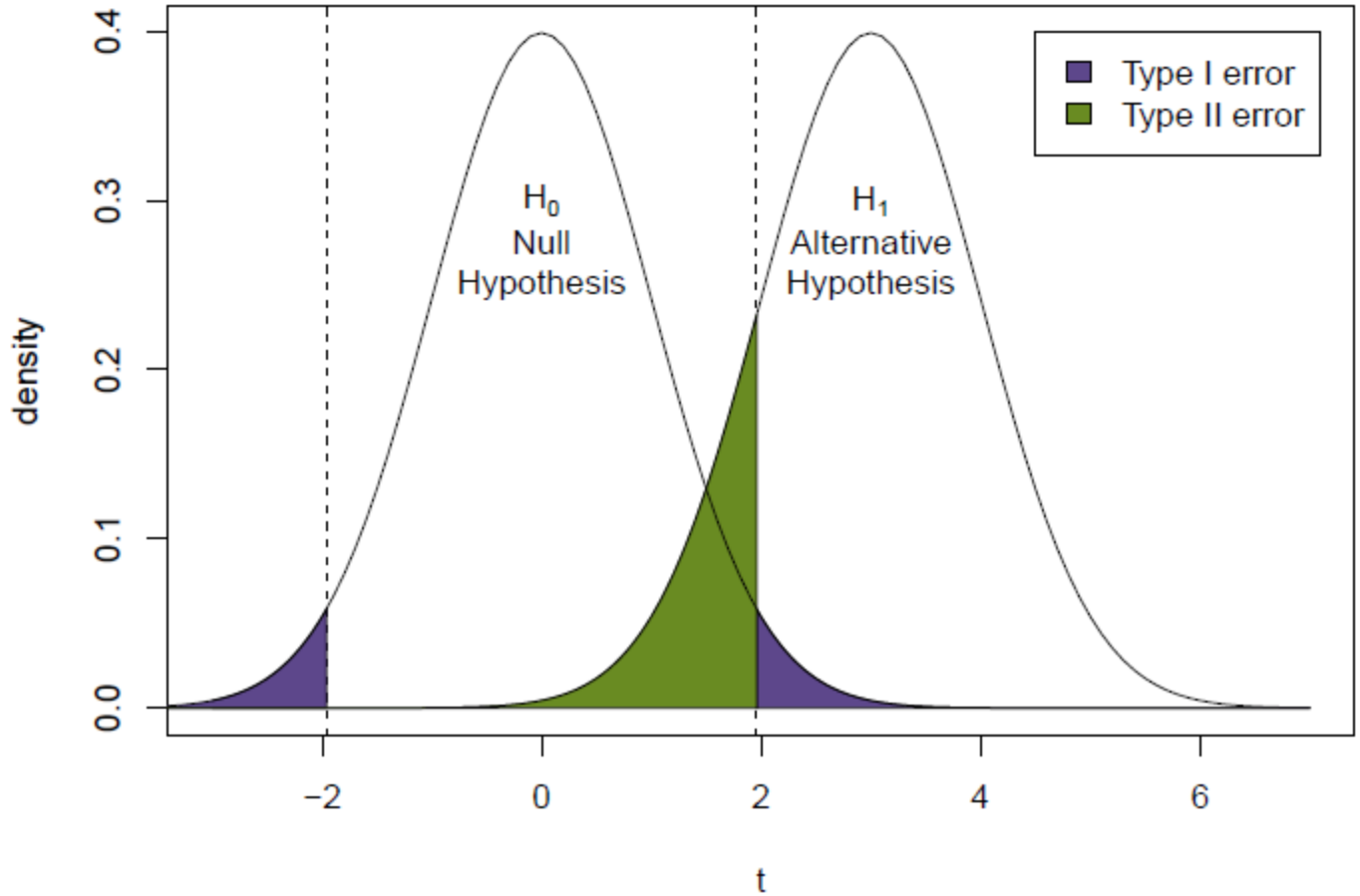
Statistical Terms: Power (1-β)

- Probability of a having a successful trial
- Probability of detecting a statistically significant difference when a true difference actually exists in the larger population
- The power of a test is a function, not a specific value, since it depends on the true difference, which is unknown

Power, α, β and Decision: How They Relate

	Decision Concerning the Null Hypothesis	
Unknown "*Truth*" of the Null Hypothesis	***Not Reject*** ***(The results for the trial groups are NOT significantly different)***	***Reject*** ***(The results for the trial groups are significantly different)***
True (*treatments are the same*)	***Correct, but undesirable condition***	***Type I Error***
False (*treatments are different*)	***Type II Error***	**Power (IDEAL)**

Correct decisions and errors in hypothesis testing



Where are α, β, & Power represented above?

What 'Exactly' is a P-value?

First, let's re-summarize the following concepts:

- **Population: Largest collection of entities for which we have an interest at a particular time**

- **Population parameters: Unknown parameters representing the theoretical value for the measure in the entire population**

- **Sample: A representative subgroup of the population (b/c we can not sample the entire population of interest)**

- **Statistic: An observed characteristic of a <u>sample</u> used to <u>estimate</u> population parameters**

The P-value

- In statistics, we draw conclusions (inferences) about the population of interest based on sample estimates

- We calculate a P-value from the SAMPLE data using appropriate statistical test (e.g. t-test)

- **P-value is conditional on the data**; probability of the observed test result (or result more extreme) given a true H0

P-Value Interpretations

- "The probability that trial results were due to chance and given a small p-value one can conclude with confidence that the results are true or real". Is this correct?
  - Answer: No, but why?
  - Opposite of 'chance' is not 'truth'…
- "The probability that the null hypothesis is TRUE." Is this correct?
  - Answer: No. First, the null is simply a statement. Second, the p-value gives us evidence (or lack thereof) against the null.

How to Correctly Interpret a P-value

- P-value is the probability (P) of observing a test result (difference, ratio, etc.) as extreme or more extreme, than that observed, if H0 were actually <u>true</u>
  - <u>Form of deductive reasoning: State with H0 and predict what we should see if H0 is true</u>

- Small p-values (i.e. very low probability) give more evidence against the null hypothesis
- Provides a summary of the degree in which the data are consistent with the null hypothesis

- *Note: This is NOT the same as 'probably of H0 true, given the observed test result'. This is the Bayesian (inductive) interpretation and assumes some prior knowledge/information.*

P-value Errors in Interpretation to Avoid

- 'A small p-value implies a high probability of observed result is true'
  - Ex: point estimate of 6.3, p-value=0.014 for a H0: difference=0. This particular H0 does not provide a probability estimate of what the true value might be but only what it is not. The confidence interval help to estimate the range of the true value (discussed shortly)
- 'A non-significant p-value is good evidence that H0 is true'
  - **Absence of evidence is not evidence of absence**
- 'P-values are interpreted the same regardless of study design'
  - Again, the pre-specified hypothesis drives the interpretation

Statistical Significance

- Comparing calculated p-value against specified type I error
- If $P < \alpha$, we declare "statistical significance" and reject the null hypothesis, which means concluding the null is not true
- The statistical analysis of a trial leads to a decision to *reject* or *fail to reject* the pre-specified *null hypothesis*

- Failing to reject the null hypothesis is *NOT* the same as accepting the null hypothesis
  - **Absence of evidence is not evidence of absence**

- *P* value & statistical significance provide little information about the magnitude of the effect being studied
  - **Large sample can yield a low p-value for a small effect size**

Confidence Interval

- A set of two values (upper and lower bound) calculated from the data around the treatment effect/estimate (usually the mean); provides a range for the TRUE value

- Provides *'level of confidence'* that the true parameter (that is being estimated from the sample such as difference in average HAMD scores between treatment groups) is between the lower and upper bounds of the confidence interval

Confidence Interval

- Width provides a level of precision around point estimate
- The CI is a direct function of sample size, the more precise the estimate the narrower the CI
- Statistical significance demonstrated if zero (difference) or one (ratio) is excluded from the CI
  - **Example 1: 95% CI of 0.218 and 0.470 provides us with 95% confidence that the TRUE mean change from baseline in HAM-D falls between 0.218 and 0.470**
  - **Example 2: 95% CI ranging from -1.4 to -3.5 around point estimate (difference in proportion of patients meeting clinically relevant change from baseline in HAM-D) of -2.4**

Interpreting the Confidence Interval

- Developed as a way to deal with the uncertainly inherent in the results obtained from the randomly selected sample from the entire population of possible samples.

- A 95% confidence interval is a set of parameter values formed by procedure, which is repeated many times, will contain the true population parameter 95% of the time.
  - EX: A 95% CI of -5 to -2 around the point estimate of -3.5 tells us that, on average, 95 of the 95% CIs calculated would contain the true value (and 5 would not).

- The X% CI means that we would expect X% of the interval estimates (from repeat experiments/sampling) to include the <u>true population parameter</u>.

Interpreting the Confidence Interval

- The X% CI does NOT mean there is a X% chance that the population parameter is between the upper and lower bounds.
  - This is different and is the called the probability interval
  - A Bayesian approach that assumes prior assumptions or pre-test probability
  - Example: A trial to measure rate of MI in 10 patients of which all 10 experience an MI. The 95% CI would be 0% to 31%. The 95% probability interval (based on some prior information that the MI rate is 20%) would be 4% to 28%. A difference b/c of the assumed information.
- So, take careful note in how the 95% CI is calculated and interpreted

Fundamental Statistical Tests

Continuous Measure: Testing of a Single Mean

- One random sample drawn from population of interest
- Single distribution or paired observations
- Hypotheses:
  - H_0: mean change from BL = 0 (or other pre-specified amount)
  - H_1: mean change from BL ≠ 0

- Variance known: One sample Z-test
- Variance unknown: One sample t-test
- Nonparametric tests (i.e. when data are small (usually n<30))
  - Wilcoxon signed-rank test

Continuous Measure: Comparison of Two Means

- Example: Comparing mean change from BL between samples
- Hypotheses:
  - H_0: $mean_1 - mean_2 = 0$
  - H_1: $mean_1 - mean_2 \neq 0$

- Parametric Tests (test of sample means)
  - Unpaired/independent data: Two-sample t-test
  - Paired/dependent data: Paired t-test (e.g. BL measure, post-treatment measure in same individuals; measure is change)

Continuous Measure: Comparison of Two Medians

- Example: Comparing medians between groups when sample size is small (usually <30)
- Hypotheses:

$$H_0: \text{median}_1 - \text{median}_2 = 0$$

$$H_1: \text{median}_1 - \text{median}_2 \neq 0$$

or

$$H_0: \text{median}_1 - \text{median}_2 \leq 0$$

$$H_1: \text{median}_1 - \text{median}_2 > 0$$

- Wilcoxon signed-rank test: paired samples
- Wilcoxon rank-sum test: independent samples

Comparison of ≥ 3 Means

- Parametric test: Analysis of Variance (ANOVA)
  - Generalization of the t-test
- Nonparametric test: Kruskal-Wallis test
  - Generalization of Wilcoxon rank-sum test

Testing of One Proportion

- Z-test
- Exact binomial calculation

Comparison of Two Proportions

Parametric test: Z-test

- Example: comparing proportion of responders to treatment
- Assumptions
  - Independent samples
  - Samples from binomial distributions, known SD (used t-test otherwise)
- Hypotheses

$$H_0: \pi_1 - \pi_2 = 0$$

$$H_1: \pi_1 - \pi_2 \neq 0$$

or

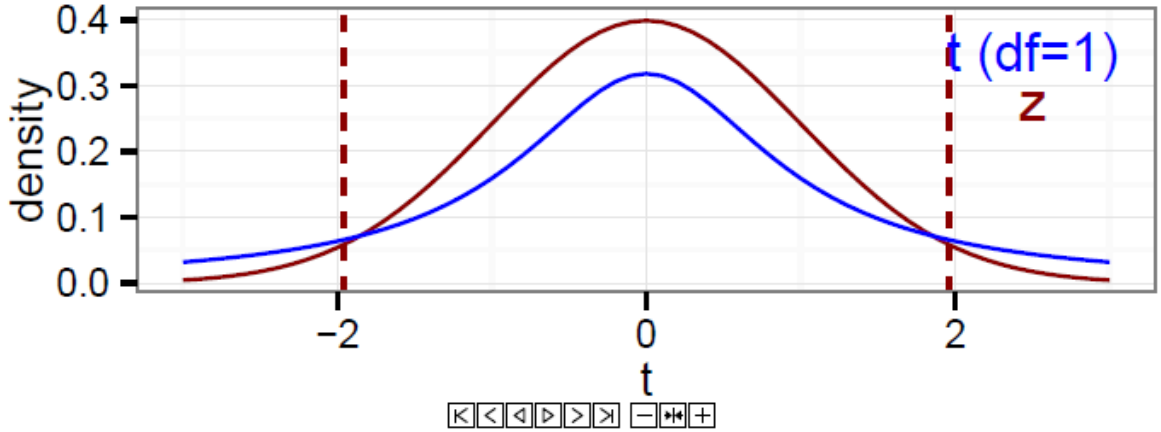
$$H_0: \pi_1 - \pi_2 \leq 0$$

$$H_1: \pi_1 - \pi_2 > 0$$

- Alternative: Fishers exact if number of subject*proportion is ≤ 5

T-Test when SD is Unknown

z- vs t-distribution



- t-distribution is symmetric, like z
- Assumes a different shape for each sample size (n), via degrees of freedom (df) $n-1$
- As $(n-1) \to \infty$, t-distribution gets closer to the z
- "Heavier tails" for smaller n.
- t-distribution provides more conservative rejection region (dashed lines)

What about Chi-Square Test?

- Test to test for an association between treatment (exposure) and outcome in at least two groups (with at least 2 outcomes)
  - Different than Z-test for proportions
  - Just keep in mind that this is used for proportions

Chi-square test for any r by c data

To test the hypothesis that

Ho: The row variable and column variable are independent vs

Ha: The row variable and column variable are associated.

The Chi-square test statistic is calculated just as it was earlier for the 2 x 2 table.

$$\chi^2(df) = \sum_{\substack{all \\ cells}} \frac{(O-E)^2}{E}$$

$$df = (r-1)(c-1)$$

Comparison of ≥ 3 Proportions

- Parametric test:
  - Chi-square with k-1 degrees of freedom, where k = number of independent binomial proportions
  - Generalization of z-test
- Nonparametric test: Exact test
  - Generalization of Fisher exact test

Summary of Statistical Tests for continuous Outcomes

Groups	Purpose	Assumptions	Statistical Test
One sample	Compare mean to historical or hypothetical mean	Standard deviation is known or sample size is large and data is approx normal	One-sample z-test
One sample	Compare mean to historical or hypothetical mean	Standard deviation unknown; approximate normality	One sample t-test
Two independent samples	Compare means	Approximate normality	Two-sample t-test
2+ independent samples	Compare means (any overall difference)	Approximate normality	ANOVA
Matched pairs	Test if mean of differences is not equal to zero	Approximate normality	Paired t-test
2+ matched sets	Test if mean of differences is not equal to zero	Approximate normality	Repeated measures ANOVA

Summary of Statistical Tests for Binary Outcomes

Groups	Purpose	Assumptions	Statistical Test
One sample	Compare proportion to historical or hypothetical proportion	Large N	One-sample z-test (exact binomial for small N)
2+ independent samples	Compare proportions or test for association	sufficiently large expected frequencies	Chi-square test
2+ independent samples	Compare proportions or test for association	expected frequencies are small	Fisher's Exact test
Matched pairs	Compare proportions or test for association		McNemar's test

Risk and Odds

- Risk: Probability of an event occurring: event divided by total number at risk of having the event
  - Example: 20 events in 100 leads to a risk of 20/100 or 0.20
- Relative Risk: Ratio of risks
- Risk Difference: Difference in risks
- Odds: Probability of an event occurring divided by 1-probability of event occurring
- Example: 20 events in 100 leads to an odds of 20/80= 0.25

Correlation

- Correlation (r)– measures the nature and strength of linear association between two variables at a time

- Regression – equation that best describes relationship between variables

Regression Methods

- Linear regression: modeling a continuous outcome while including patient-level characteristics
- Logistic regression: modeling a dichotomous/binomial outcome while including patient-level characteristics
- Overarching goal: Estimate amount of model variance is solely due to independent variable(s) such as treatment

Simple Linear Regression Model

The value of any observation (Y_i) is modeled as:

$$Y_i = \beta_0 + \beta_1 X_i + e_i$$

The value of Y for observation (i) is made up of

- An intercept (β_0)
- A sloping function of the observation's value on variable X (β_1)
- An error term (*e*), the deviation from the line
- Ex: Modeling treatment on change in HAM-D at day 30

Multivariable Linear Regression

The full multiple regression model is:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + e_i$$

A slope (β) is estimated for each X variable

Single dependent variable Y

Error e_ι is normally distributed w/ mean 0 and variance σ^2

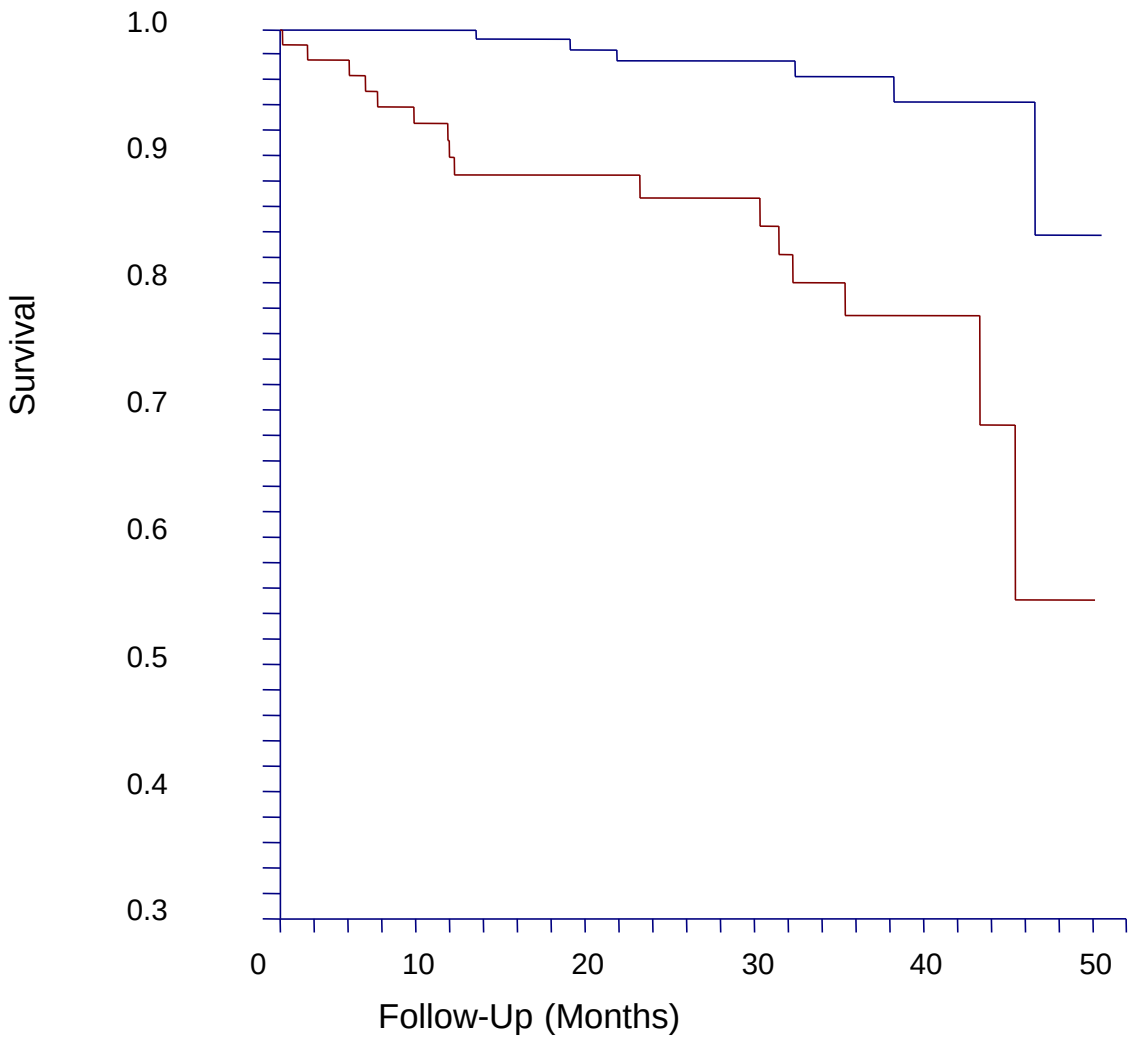
EX: Treatment effect on change from baseline in HAM-D with age, sex, baseline treatment status (y/n) as predictors

Survival Analysis: Time to Event

- Like binomial proportions, have 0-1 variable (success/failure, alive/dead)
- In general, event is not observed for all subjects, and not all subjects are followed for the same length of time; there is *censoring* – subject who reaches end of study or is lost to follow-up without experiencing event is considered to be censored

Survival Analysis: Time to Event

- Data often presented as proportion surviving (Kaplan-Meier)
  - Show proportion who have not experienced event (y-axis) as function of time (x-axis)
  - Subjects not reaching endpoint are censored at end of follow-up or last observation (right censor)


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Survival
0
10
20
30
40
50
Follow-Up (Months)

Survival Analysis: Time to Event

Comparison of groups

- Log-rank test compares median survival times
  - Peto
  - Mantel
- Proportional hazards (Cox) regression
  - Hazard: instantaneous risk
  - Model assumes ratio of hazards is constant (proportional hazards assumption)
  - Model incorporating predictor variables