

Bio-statistical Analysis of Research Data

March 18th and March 25th, 2016

Kris Attwood, PhD
Department of Biostatistics & Bioinformatics
Roswell Park Cancer Institute

Outline

- Biostatistics in Research
- Basic Concepts
- Common Analyses
- Statistics for Grants and Protocols
- Limitations

Biostatistics in Research

- **Statistics**

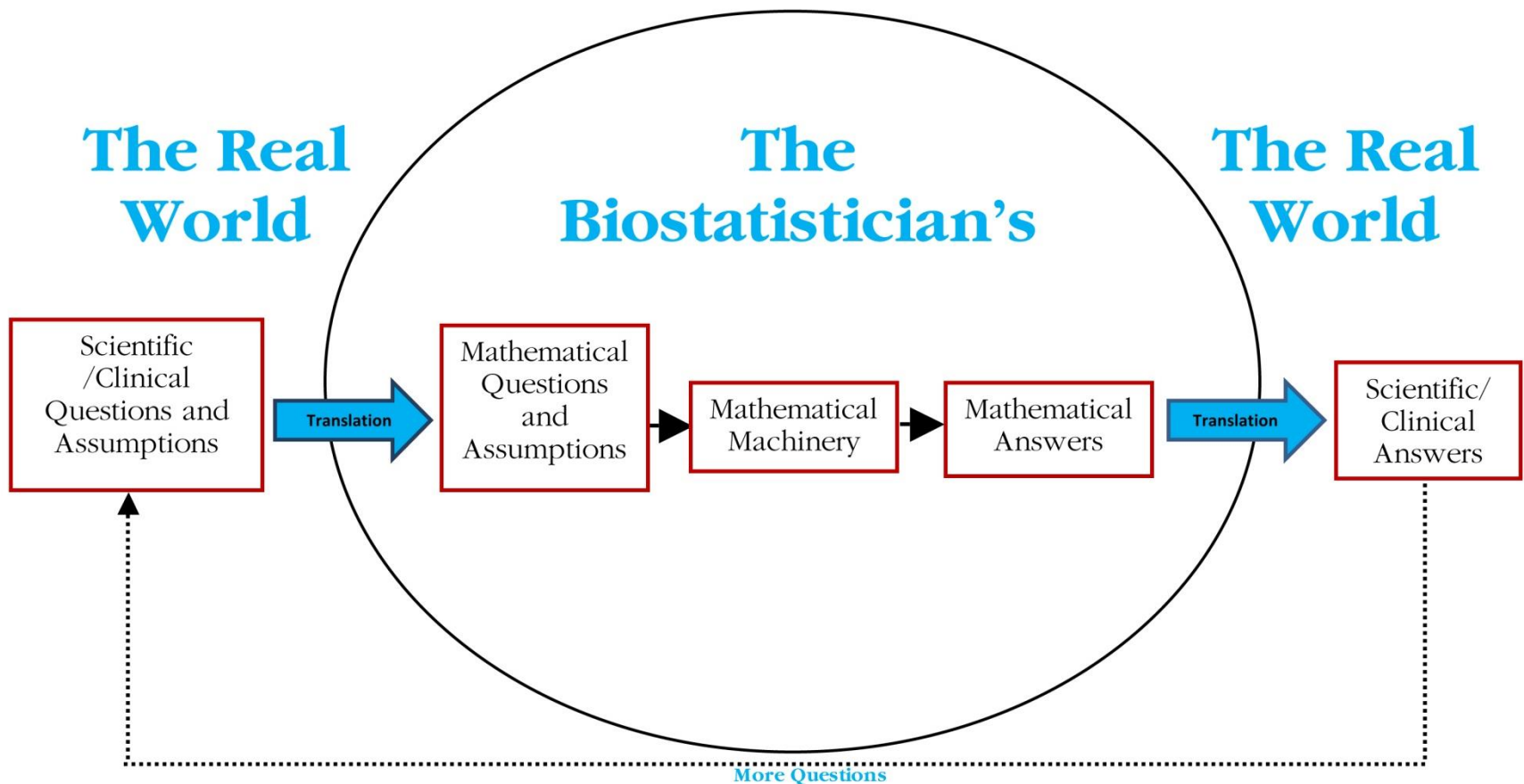
- Formal Definition

- “...a collection of mathematical methods for organizing, summarizing, analyzing and drawing conclusions based on data gathered in a study.”

- Practical Application

- Using data to derive conclusions about clinical research questions.

Biostatistics in Research



Comparative Analysis

- Comparison of variables between two or more groups (treatments/conditions)
 - Any difference?
 - Is there a particular order?
- Example: Comparing expression between groups.
 - The purpose of the study is to determine whether gene expression is different between responders and non-responders
- Example: Comparing rates between groups.
 - Is there a difference in the complete response rate between two treatment groups?

Correlative Analysis

- Evaluating and quantifying the relationship between two (or more) variables
 - Evaluating
 - Direction
 - Magnitude
 - Quantifying
 - Predictive models
- Example: Correlation between two (or more) measures.
 - The study goal is to quantify the relationship between mTOR and EGFR expression in solid-organ transplant patients.

Time-series Analysis

- Evaluating the behavior of a variable over time.
 - Example: Evaluating expression and time
 - The study goal is to determine how PSA expression changes after castration.
- Comparison of variables between groups over time or with repeated measures.
 - Example: Comparing tumor growth between groups
 - The study goal is to compare tumor volume in mice observed under 3 different conditions.

Survival Analysis

- Comparison of *time-to-event* variables between different treatment/condition cohorts.
- Evaluating the relationship between a variable and the *time-to-an-event*
 - Issue: Do not observe the “event” for all subjects
- Example: Survival outcomes with HIPEC
 - The study goal is to examine clinical and surgical factors associated with survival in HIPEC treated patients.

Basic Concepts

- Who are we studying?
 - **Population**
 - All individuals under investigation
 - The research question applies to this theoretical group
 - **Ex.** Everyone who would use a given treatment.
 - **Ex.** Everyone who has a particular condition.
 - **Sample**
 - The individuals actually used to obtain data
- Statistics vs. Parameters
 - **Statistics** – values that summarize a sample characteristic
 - **Parameter** – values that summarize a population characteristic

Basic Concepts - Data

- Types of Data
 - **Quantitative** – measurements or counts
 - Discrete or Continuous
 - **Qualitative** – attributes or labels
 - Discrete

Basic Concepts - Data

- Levels of Measure
 - The data we obtained are simply numerical representations (**measurements**) of a characteristics.
 - Levels:
 - **Nominal** (*lowest*) – labels with no order
 - **Ordinal** – ordered data with no consistent intervals
 - **Interval** – ordered data and consistent intervals, but *no true zero*
 - **Ratio** (*highest*) – ordered data, consistent intervals and true zero

Basic Concepts – Descriptive Statistics

- Purpose

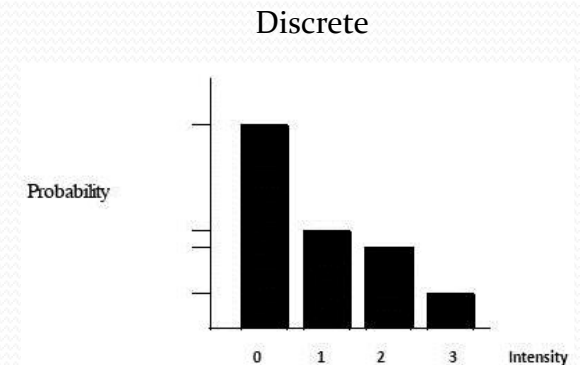
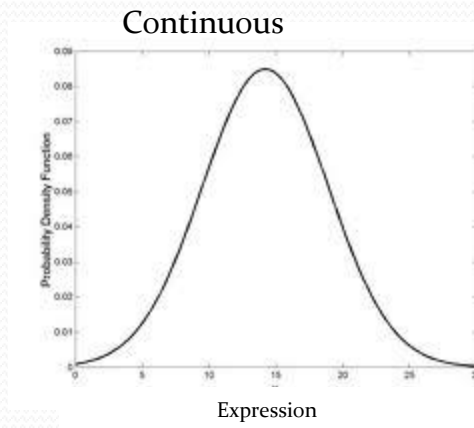
- *What story does the data from your study tell?*

- **Distribution**

- Description of the possible values for a variable and how often they occur.

- Components:

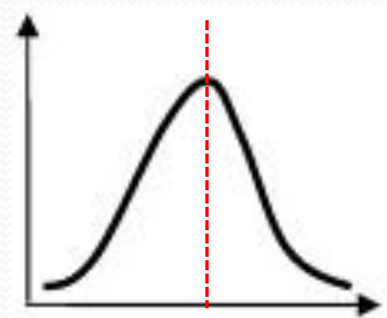
- Center
- Spread
- Shape



Basic Concepts – Descriptive Statistics

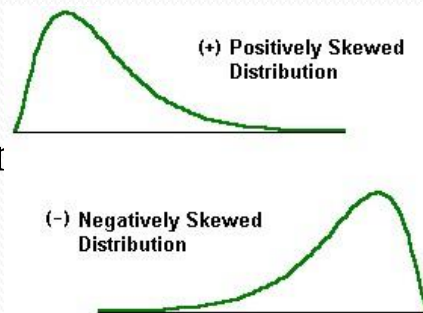
- Shapes of a Distribution

- **Symmetric**



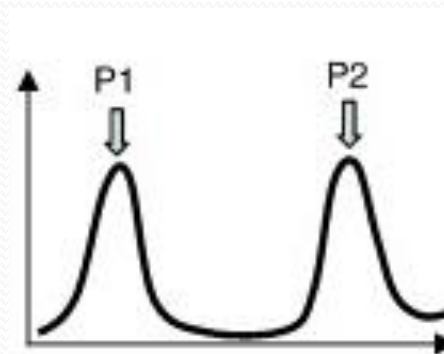
- **Skewed**

- Tail contains **outlier**



- **Bimodal**

- Two populations are mixed together



Basic Concepts – Descriptive Statistics

- Measures of Center
 - *Describing the typical or expected value*

- Statistics:

- **Mean** – the average value

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

- **Median** – the middle value
 - 50% above and 50% below

- **Mode** – the most frequent value

Basic Concepts – Descriptive Statistics

- Measures of Center
 - Example Data: Consider the following expression levels
 - **Data: 0, 2, 5, 8, 10**
 - Mean = $\frac{0 + 2 + 5 + 8 + 10}{5} = \frac{25}{5} = 5$
 - Median = 5
 - Mode = All values
 - *How do **outliers** effect these measures?*
 - *What is the last observation was 100?*

Basic Concepts – Descriptive Statistics

- Measures of Center
 - Which do we use?
 - **Mean**
 - Interval/Ratio Data and Symmetric Distribution
 - Interval/Ratio/Ordinal Data and Large Samples
 - **Median**
 - Ordinal Data
 - Interval/Ratio Data and Skewed Distribution
 - **Mode**
 - Nominal Data

Basic Concepts – Descriptive Statistics

- Measures of Variability
 - Are the observations homogeneous or heterogeneous?
- Statistics:
 - **Range** – difference between the smallest and largest value
 - **Standard Deviation** – similar to the “average” deviation from the mean
 - Deviation = difference between an observation and the mean
 - **Coefficient of Variation** – standard deviation divided by the mean
 - Accounts for the magnitude of the data
 - **IQR** – difference between the 75th and 25th percentiles
 - Spread of the middle 50%

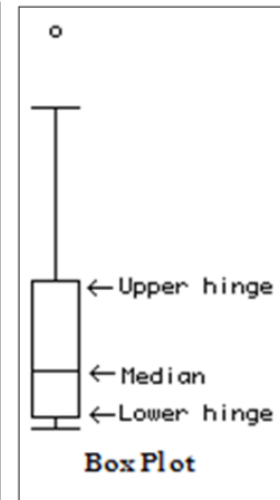
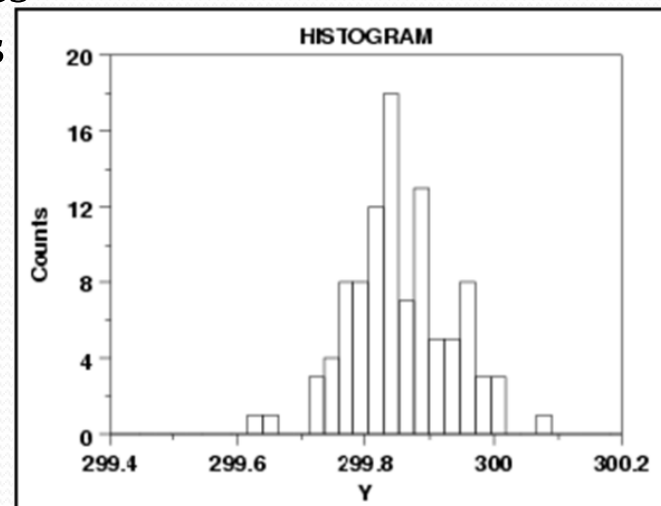
Basic Concepts – Descriptive Statistics

- Measures of Center
 - Which do we use?
 - **Mean → Standard Deviation**
 - The coefficient of variation can be used when comparing groups
 - The **standard error** of the mean is sometimes reported as well
 - **Median → IQR**

Basic Concepts – Descriptive Statistics

- Graphical Summaries
 - Exploring the Distribution
 - Histograms

- Data are lumped into *classes*
- The bar height corresponds to class frequency



- **Box Plots**

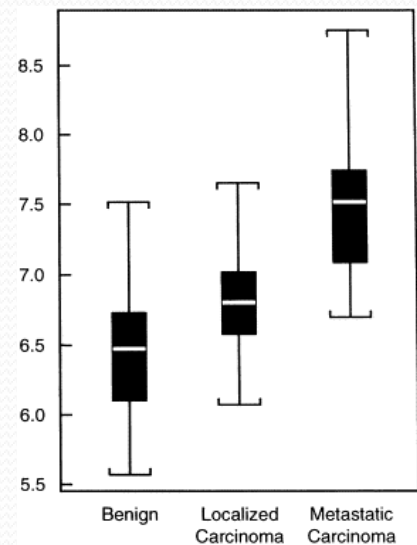
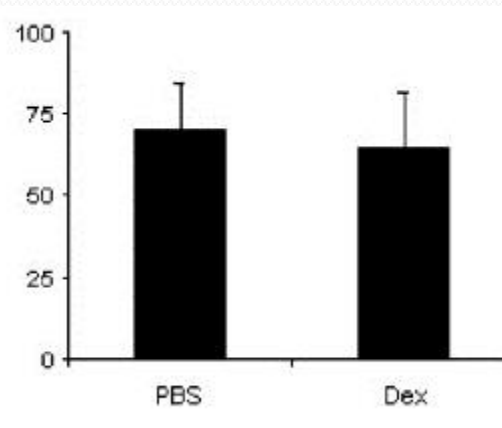
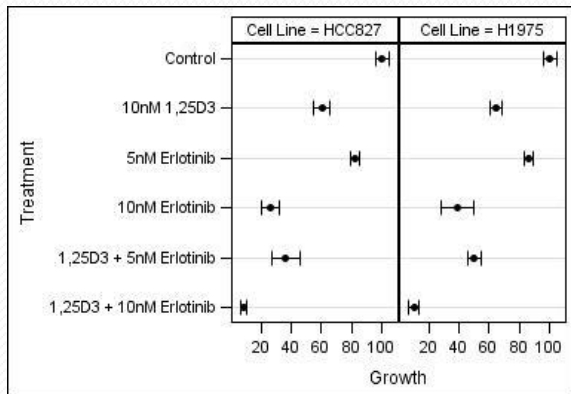
- 5-point summary
 - Minimum, 25th percentile, Median, 75th percentile, and Maximum
- Identifies *statistical outliers*
 - Outside 1.5 IQR's of the 25th or 75th percentiles

Basic Concepts – Descriptive Statistics

- **Comparisons**

- **Box Plots**

- Plot several treatment or condition cohorts on the same axis



- **Mean Plots**

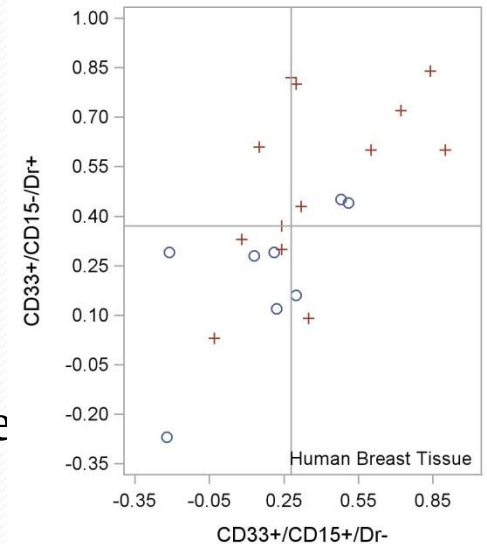
- Plot the mean for each cohort as a dot or a bar
 - Generally includes an error bar
 - 1 standard deviation or standard error

Basic Concepts – Descriptive Statistics

- **Exploring Relationships**

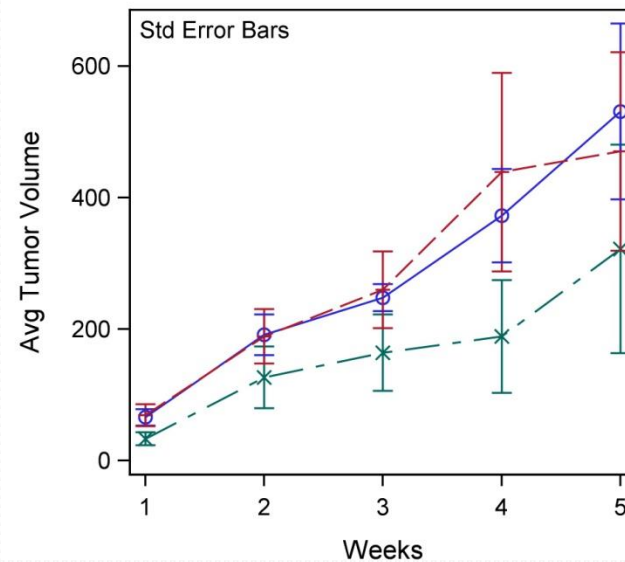
- **Scatter Plots**

- Data are treated as paired observations
 - Each dot corresponds to an observation
 - X-axis = variable 1 (independent variable)
 - Y-axis = variable 2 (dependent variable)



- **Time Series Plots**

- Data are plotted over time
 - X-axis = time
 - Y-axis = mean value

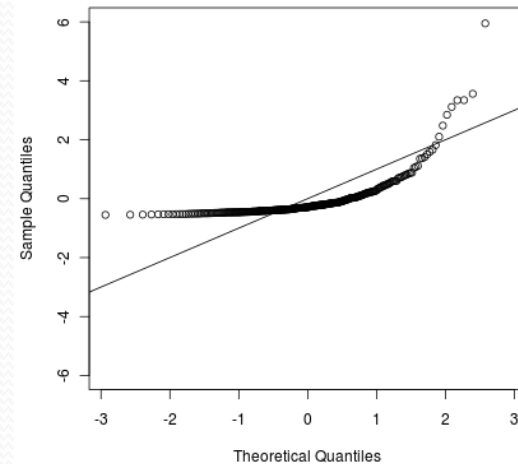
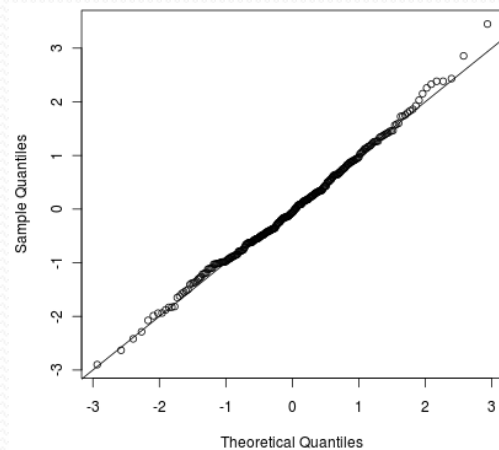


Basic Concepts – Descriptive Statistics

- **Assessing Normality**

- **QQ Plots**

- Compares the observed percentiles to the expected percentiles



- If the data are approximately *normal*, then the graph should follow the 45° diagonal
 - If the data are not normal, a **transformation** may be useful

Basic Concepts – Descriptive Statistics

- Measures of Relative Position

- **Percentiles**

- The K^{th} percentile (P_K) is the value such that $k\%$ of observations are less than or equal to that value.

- Example:

- Based on a recent study, the 75th percentile for mTOR expression in patients with pancreatic tumors was estimated to be **6**.
 - Therefore, **75%** of patients with pancreatic tumors have an mTOR expression of **6 or less**.

Basic Concepts – Confidence Intervals

- What are confidence intervals?
 - Confidence intervals provide inferential estimates of population characteristics based on sample data
- Utility
 - Statistical
 - Probabilistic interval estimate of a population parameter
 - A 95% confidence interval implies that if you repeated this experiment 100 times and calculated 100 confidence intervals, then 95 of them would contain the true parameter
 - Practical
 - Range of possible values for our parameter

Basic Concepts – Confidence Intervals

- Confidence intervals can generally be constructed for any parameter
 - Closed Form:
 - Sample Statistic \pm (Standard Score) \cdot (Standard Error)
 - The confidence level comes in through the Standard Score, which is based on the distribution of your statistic
 - Ex: Confidence interval for the mean: $\bar{x} \pm T_C \cdot (s_{\bar{x}})$
 - Bootstrapped:
 - Using bootstrap re-sampling, you can get the “exact” distribution of a statistic

Basic Concepts – Hypothesis Tests

- What is a hypothesis test?
 - A hypothesis test is a statistical method that uses data to decide between two competing hypotheses
 - Decision making tool
- Almost any research question can be boiled down into a hypothesis test
 - Later in the talk we'll look at some examples

Basic Concepts – Hypothesis Tests

- General Framework
 - Identify the *Hypotheses*
 - Specifically interested in the alternative
 - Calculate the appropriate *test statistic*
 - This is a **standardized score** based on your data
 - Has a **known distribution**
 - Calculate the corresponding *p-value*
 - Is this a one or two tailed test?
 - Make a decision
 - What is the clinical significance?



Use
statistical
software

Basic Concepts – Hypothesis Tests

- Hypotheses
 - **Null Hypothesis**
 - Hypothesis of equality
 - Prior belief
 - In the hypothesis test, we assume this to be true
 - **Alternative Hypothesis**
 - The hypothesis of change
 - Researchers belief
 - Try to disprove the null in favor of the alternative
- One or Two sided?

Basic Concepts – Hypothesis Tests

- Test Statistic

- A standardized measure of the difference between what you *observed* and what is *expected* under the null hypothesis

The diagram shows the formula for the test statistic $T = \frac{\bar{x} - \mu}{S_{\bar{x}}}$ enclosed in a blue rectangular box. The numerator $\bar{x} - \mu$ is circled in yellow, and an arrow points from this circle to the text "Difference from observed sample mean and expected population mean". The denominator $S_{\bar{x}}$ is circled in green, and an arrow points from this circle to the text "Natural error in a sample".

- Based on corresponding sample statistics
 - Ex. If you are making inferences about the population mean, then your test statistic is based on the sample mean

Basic Concepts – Hypothesis Tests

- P-value
 - The *probability* of getting you test statistic or something more in *favor of the alternative*, if the null hypothesis were true
 - Smaller p-values favor the alternative hypothesis
 - One- or Two- tailed
 - Obtained using:
 - Distribution of the test statistic
 - Bootstrap methods
- Decisions:
 - Compare the p-value to the *significance level*
 - If $p\text{-value} \leq \text{significance level}$ then Reject the Null Hypothesis

Basic Concepts – Hypothesis Tests

- Errors
 - **Type I** – Reject the Null when it is True
 - **Type II** – Fail to reject the Null when it is False

Basic Concepts – Hypothesis Tests

- Examples:

- We conduct a hypothesis test on a new drug, where the alternative hypothesis is that the toxicity rate is less than 30%
 - $H_A: TR < 30\%$
 - A type I error would lead to acceptance and further study on an unsafe drug
- We conduct a hypothesis test on a new drug, where the alternative hypothesis is that the response rate is greater than 75%
 - $H_A: RR > 75\%$
 - A type II error leads to the missed opportunity in developing a successful tx

Basic Concepts – Hypothesis Tests

- **Significance Level**

- The maximum allowed type I error
 - Pre-specified
 - General values: 0.01, 0.05, and 0.10

- **Power**

- Probability of rejecting the null hypothesis if the alternative is true
 - Can we detect a significant shift or difference?
 - Generally look for a power $> 70\%$

Basic Concepts – Types of Analysis

- Parametric vs. Non-Parametric
 - Questions:
 - What type (and level) of data do you have?
 - What assumptions do you want to make about the distribution of the data?
 - **Parametric Analysis**
 - Assumes the data follows given distribution
 - Ex: the most common is the normal distribution
 - **Non-Parametric Analysis**
 - No set distributional assumptions are required

Common Analysis

- Comparative
 - 2 Groups
 - T-test
 - Wilcoxon rank sum
 - 3+ Groups
 - ANOVA
 - Kruskal Wallis
 - Categorical
 - Chi-Square
 - Fisher's Exact Test
- Repeated Measures
 - Repeated Measures ANOVA
 - Friedman ANOVA
- Correlative
 - Correlation Coefficients
 - Regression
 - Mixed model Regression
- Survival Analysis
 - Kaplan-Meier
 - Cox Regression

Comparing Groups: Continuous Data

- General Purpose:
 - Are there any differences in the values between 2 or more groups?
 - Example: Is there an **association** between *mTOR* expression and immunosuppression?
- Comparing Independent Samples
 - Independent Samples
 - Two or more separate groups of subjects
 - Parametric vs. Non-parametric
 - **Parametric**
 - Assumes the data are normal
 - Interval or Ratio Data
 - **Non-Parametric**
 - No distributional assumption
 - Ordinal, Interval or Ratio data

Comparing Groups: Continuous Data

- T-test

- Parametric Test for 2 Samples
 - Requires approximately normal data

- Hypotheses:

- With respect to the difference in **mean** values

- Alternatives:

• $H_A: \mu_1 \neq \mu_2$	\leftrightarrow	$H_A: \mu_1 - \mu_2 \neq 0$
• $H_A: \mu_1 > \mu_2$	\leftrightarrow	$H_A: \mu_1 - \mu_2 > 0$
• $H_A: \mu_1 < \mu_2$	\leftrightarrow	$H_A: \mu_1 - \mu_2 < 0$

- Test Statistic:

- Compares the observed difference to natural variability

$$T = \frac{\bar{x}_1 - \bar{x}_2}{S_{x_1 - x_2}}$$

Comparing Groups: Continuous Data

- **Wilcoxon rank sum**

- Non-parametric Test for 2 Samples
 - Makes no distributional assumptions
 - Useful for *ordinal* data

- Hypotheses:

- With respect to the difference in **median** values
- Alternatives:

• $H_A: M_1 \neq M_2$	\leftrightarrow	$H_A: M_1 - M_2 \neq 0$
• $H_A: M_1 > M_2$	\leftrightarrow	$H_A: M_1 - M_2 > 0$
• $H_A: M_1 < M_2$	\leftrightarrow	$H_A: M_1 - M_2 < 0$

- Test Statistic:

- Compares the observed ranks to the expected ranks
- Based on ranks, not on the actual data – thus only requires the data has *order*

Example: Tumor Volume Reduction

- Comparison of Tumor Volume
 - The study goal is to compare the effectiveness of two treatments in reducing tumor volume. Two cohorts of 20 mice each are treated with Tx-A or Tx-B and the reduction in tumor volume is recorded after 2 weeks.
- *Study Design:*
 - Data (volume reduction) is collected on 40 mice (20 in each treatment)
- *Question:*
 - Is there a difference in reduction between these two groups?

Example: Tumor Volume Reduction

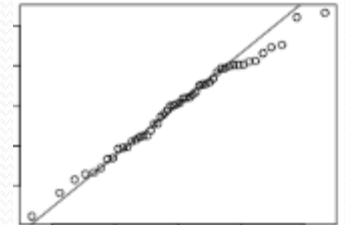
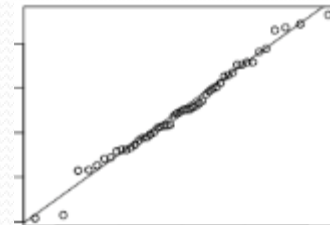
- Data Analysis

- Summary Statistics:

- *What type of data do we have?*
 - We would consider this as *interval/ratio* data

- *Is the data normally distributed?*

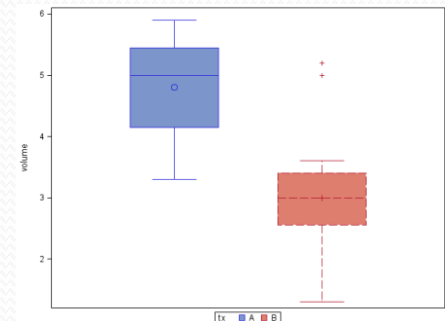
- Use QQ plots



- *Describe the center and variability.*

- Since the data appears normal, we use the mean and standard deviation

tx	N	Mean	Std Dev	Std Err
A	20	4.8050	0.8457	0.1891
B	20	2.9900	0.9380	0.2097
Diff (1-2)		1.8150	0.8931	0.2824



Example: Tumor Volume Reduction

- Hypothesis Test:
 - *What type of test?* **Normal data** → **T-test**
 - Hypothesis:
 - $H_0: \mu_A = \mu_B$
 - $H_A: \mu_A \neq \mu_B$
 - Test Statistic and P-value:
 - $T = 6.43$
 - P-value **<0.0001**

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	38	6.43	< .0001
Satterthwaite	Unequal	37.599	6.43	< .0001

- Conclusion:
 - *Statistical:* Reject the null hypothesis
 - *Clinical:* There is a significant difference in the reduction of tumor volume between treatments A and B

Comparing Groups: Continuous Data

- ANOVA

- Parametric Test for 3+ Samples
 - Requires approximately normal data for each sample
- Hypotheses:
 - With respect to the difference in **mean** values
 - H_0 : All μ 's are equal
 - H_A : At least two μ 's are different
- Test Statistic:
 - Compares the variability between groups (group differences) to natural variability between observations

Comparing Groups: Continuous Data

- **Kruskal-Wallis**

- Non-parametric Test for 3+ Samples
 - Makes no distributional assumptions
- Hypotheses:
 - With respect to the difference in **median** values
 - H_0 : All M's are equal
 - H_A : At least two M's are different
- Test Statistic:
 - Compares the observed ranks to the expected ranks

Comparing Groups: Multiple Tests

- The ANOVA and Kruskal-Wallis tests can only demonstrate that there is a difference between the groups
 - It does not tell you which groups are different
 - Need to conduct a series of pair-wise tests
- Issue:
 - Each test has an error rate associated with it
 - The more test you conduct, the higher your *overall error rate*
 - Ex. If you conduct two hypothesis tests at a significance level of 0.05, then your overall chance of an error is $0.05 + 0.05 = 0.10$.
- Adjustments need to be made for multiple tests
 - Bonferonni
 - Tukey

Comparing Groups: Continuous Data

- What about *Repeated Measures* or *Paired Data*?
 - Study Designs
 - **Paired Design**
 - Two samples are compared
 - Each subject in sample 1 is linked to a subject in sample 2
 - **Repeated Measures Design**
 - Multiple samples are compared
 - Each subject has multiple observations
 - Observed in each sample
 - Issues
 - Observations from the same subject are inherently correlated

Comparing Groups: Continuous Data

- **Paired Data**
 - Scenario
 - Case-Control Study
 - Ex: Tumor vs. Healthy tissue
 - Before and After Study
 - Ex: Expression before and after treatment
 - Research Questions
 - With respect to the mean or median *difference*
 - *Difference Scores* as the data
 - Calculate the difference (case-control or before-after) for each subject
 - Tests:
 - **Paired T-test**
 - **Wilcoxon signed rank test**

Comparing Groups: Continuous Data

- **Repeated Measures Data**
 - Scenario
 - Longitudinal Studies
 - Ex: Observe Expression levels prior to surgery, after surgery and 1 year post surgery.
 - Repeated Designs
 - Ex: Patients are evaluated on each of 3 different blood pressure medications
 - Research Questions
 - Same as ANOVA type hypothesis
 - Each subject has an observation in each sample (hence repeated measures)
 - Tests:
 - **Repeated Measures ANOVA**
 - **Friedman ANOVA**

Comparing Groups: Categorical Data

- General Purpose:
 - Evaluating the association between *categorical* variables and cohorts
- Data Summary
 - Nominal data are generally presented as percentages or in tabular form.
- Example: Treatment Type and HER2 Status

- Tables:

Treatment / HER2	+	-
A	10	40
B	15	35

- Percentages:
 - 20% of patients on treatment A were HER2+
 - 30% of patients on treatment B were HER2+

Comparing Groups: Categorical Data

- Independent Samples
 - **Fisher's Exact Test:** 2x2 Tables
 - Is there an association between two dichotomous variables?
 - Most Common Use:
 - Compare proportions between two cohorts
 - **Chi-Square Test:** IxJ Tables
 - Is there an association between two categorical variables?
 - No restriction on the number of *levels*
 - Tests whether the column and row variables are independent
 - H_o : Independent
 - H_A : Dependent
- Compares *expected* cell counts to *observed* cell counts

Example: HER2 Expression

- Comparison of HER2 Expression between treatments
 - *Research Question*
 - Is there a significant difference in the rate of HER2+ (P^+) between treatments A and B?

• *Data:*

Treatment / HER2	+	-
A	10	40
B	15	35



$$\widehat{P}^+_A = 0.20$$

$$\widehat{P}^+_B = 0.30$$

• *Hypothesis Test:*

• Hypothesis

- $H_0: P^+_A = P^+_B$
- $H_A: P^+_A \neq P^+_B$

• P-value:

- Using Fisher's Exact test: $p\text{-value} = 0.093$

Comparing Groups: Categorical Data

- Dependent Samples

- Scenario

- Paired Design (Before-After or Case-Control Study) where the outcome of interest is dichotomous.

- **McNemar's Test:** 2x2 Tables

- Comparing proportions between two dependent cohorts
 - Example: Comparing expression of an antibody before and after Tx

Before/ After	+	-
+	10	10
-	5	15



$$\widehat{P}_{Pre}^{+} = 0.50$$

$$\widehat{P}_{Post}^{+} = 0.25$$

- Test: $H_0: P_{PRE}^{+} = P_{POST}^{+}$ versus $H_A: P_{PRE}^{+} \neq P_{POST}^{+}$

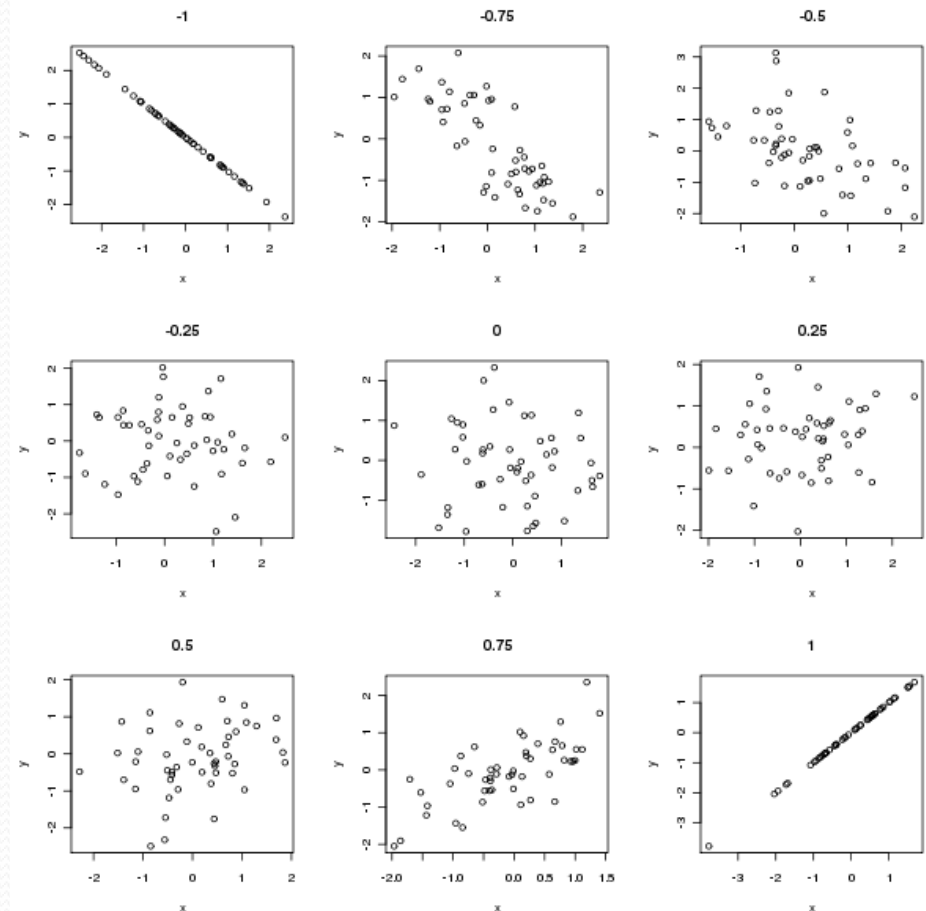
Correlation

- General Purpose:
 - Examine the relationship between two variables
 - *Direction*
 - Is the relationship positive (direct) or negative (inverse)?
 - *Magnitude*
 - How strong is this relationship?
- Typical Scenario:
 - For each subject, you record two variables and these variables are treated as “paired data”

1	mTOR	EGFR
2	4	3
3	7	3
4	2	4
5	6	3

Correlation

- Correlation Coefficients
 - Statistical measures of the direction and strength of a relationship
- Values range from -1 to 1
 - Values close to zero indicate a weak relationship



Correlation

- **Pearson Correlation Coefficient**

- Measure of the linear relationship between two interval/ratio variables
 - Only looks at linear relationships
 - Only for interval/ratio data

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

- **Spearman Rank Correlation Coefficient**

- A non-parametric measure of the relationship between two variables
 - Looks for a “monotonic” relationship between the data
 - Allows for ordinal data

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Correlation

- Clinical Relevance
 - With a large sample size, any correlation can become statistically significant
- When are these values clinically relevant?

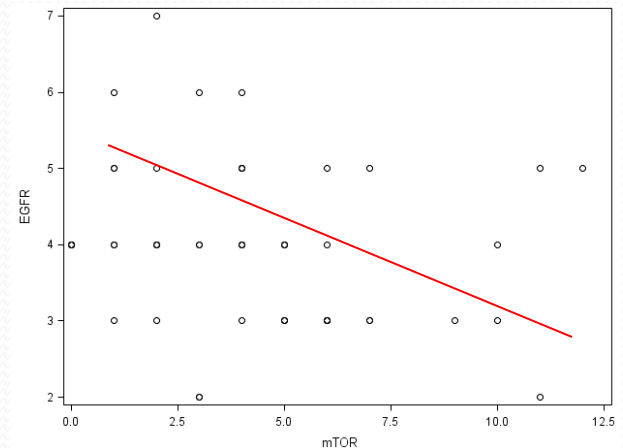
Correlation	Strength of Relationship
0 – 0.25	Little or None
0.25 – 0.50	Fair
0.50 – 0.75	Good
0.75 +	Excellent

Regression

- General Purpose:
 - Regression models can be used to:
 - Quantify the relationship between two variables
 - “For a unit increase in mTOR, how much does EGFR change?”
 - Develop prediction tools
 - “If a patient has an mTOR expression of 4, what would I expect their EGFR expression to be?”

Regression

- Simple Linear Regression
 - Model: $Y = a + bX + e$
 - Y = the response variable
 - X = predictor variable
 - a = intercept
 - b = slope
 - The effect of the predictor on the response
 - e = natural error
 - Model Fit:
 - What line would best “fit” the data?



- Method of Least Squares – generates the line with the smallest *residuals*

Regression

- Simple Linear Regression: mTOR and EGFR
 - Model EGFR as a function of mTOR
 - Response:
 - Predictor:

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	2.62615	2.62615	2.32	0.1344
Error	48	54.35385	1.13237		
Corrected Total	49	56.98000			
	Root MSE	1.06413	R-Square	0.0461	
	Dependent Mean	3.98000	Adj R-Sq	0.0262	
	Coeff Var	26.73693			
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	4.30543	0.26137	16.47	<.0001
mTOR	1	-0.07568	0.04970	-1.52	0.1344

- Model Estimate: $EGFR = 4.31 - 0.076(mTOR)$
 - *Is the model significant?*

Regression

- Simple Linear Regression Inferences
 - Test the Overall Model
 - Hypotheses:
 - H_0 : all b 's = 0
 - H_A : not all b 's = 0
 - Test Statistic:
 - **F-test** – compares the variability of the response explained by the predictor to the overall variability in the response
 - Test for an Individual Slope
 - Hypotheses:
 - H_0 : $b = 0$
 - H_A : $b \neq 0$
 - Test Statistic:
 - **T-test**

Regression

- Model Diagnostics and Transformations
 - Model Assumptions
 - We assume that:
 - The residuals are normally distributed
 - The residuals have a constant variance
 - Residuals are the differences between the predicted and observed responses
 - Diagnostic Plots:
 - QQ Plot
 - Residual Plot
 - Plots the residuals (y-axis) by the observed predictor values (x-axis)
 - Transformations

Regression

- Multivariate Linear Regression
 - Model includes multiple predictors
 - Model: $Y = a + b_1(X_1) + b_2(X_2) + \dots + b_k(X_k) + e$
 - Tests of interest:
 - Overall Model
 - F-test
 - Each Predictor
 - T-test for each predictor
 - Model Assumptions
 - Same as for a simple linear regression model

Time Series Regression

- General Scenario:
 - In some instances you may be observing subjects over time (repeated measures) and are interested in the relationship between time and some outcome.
 - Ex: The relationship between tumor volume and time for certain tx groups
- Issues:
 - The observations from a single subject are inherently correlated
 - The standard linear regression models are not able to account for this correlation

Time Series Regression

- Mixed Model: Random Subject Effect Model

- $Y_{ij} = a + bX_{ij} + \varepsilon_i + e_{ij}$

- Y_{ij} = the j^{th} response value for the i^{th} subject
 - X_{ij} = the j^{th} predictor value for the i^{th} subject
 - a = intercept
 - b = slope
 - Effect of time or another repeated predictor measure
 - e_{ij} = residual error
 - ε_i = random subject effect
 - This is where the time dependent correlation between a subject's observations is accounted for
 - You can specify several different correlation structures
 - For example, the correlation between observations at time 1 and 2 is stronger than between times 1 and 5

Time Series Regression

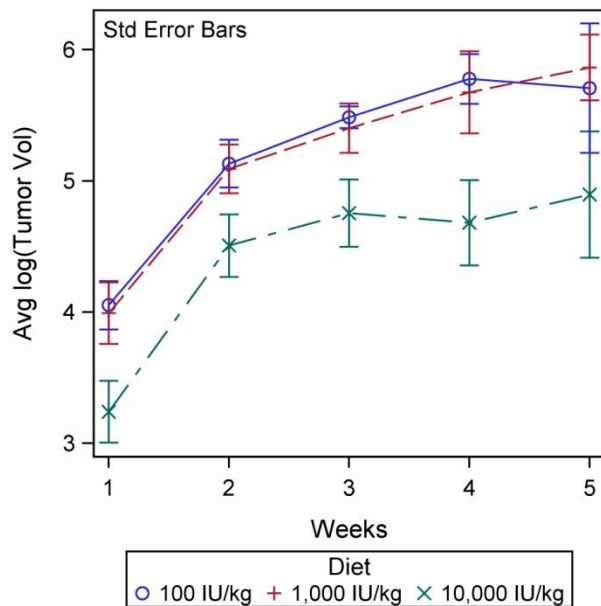
- Model Assumptions
 - The residuals are normally distributed
 - The residuals have a constant variance
- Test of Interest:
 - Overall Model
 - Model coefficients (slopes)
- **Variability associated with the ε_i**
 - This test determines whether you need to consider this mixed model approach

Time Series Regression

- More Complex Models
 - You can include in your model:
 - Other predictors
 - Other random effects
 - Interaction terms
 - In these more complex models we may be interested in testing **Combinations of Coefficients**
- Missing Data
 - You do not need the same number of observations for each subject

Time Series Regression: Example

- Tumor Growth Rate and Diet
 - A 5 week study is conducted in order to determine the effect diet on the tumor growth rate in mice. In each group, 20 mice are observed with their tumor volume assessed weekly.



The following model was considered:

$$\begin{aligned}\log(\text{Tumor volume})_{ij} = & a + b_0 \text{Time}_j \\ & + b_1 X_{1i} + b_2 X_{2i} \\ & + b_3 X_{1i} \times \text{Time}_j \\ & + b_4 X_{2i} \times \text{Time}_j \\ & + \varepsilon_i + e_{ij}\end{aligned}$$

where: $X_1 = 1$ if on 1,000 IU/kg diet
 $X_2 = 1$ if on 10,000 IU/kg diet

Time Series Regression: Example

- We use an approach similar to Least Squares to fit the model.
 - Most software programs will provide:
 - Overall Model Tests
 - *Parameter Estimates*
 - *Tests for each parameter*

Solution for Fixed Effects						
Effect	Diet	Estimate	Standard Error	DF	t Value	Pr > t
Intercept		3.5077	0.3503	4	10.01	0.0006
time		0.4094	0.08543	132	5.85	<.0001
Diet	1,000 IU/kg	0.5038	0.4767	8	1.25	0.2481
Diet	10,000 IU/kg	-0.4880	0.4729	8	-1.03	0.3322
Diet	100 IU/kg	0
time*Diet	1,000 IU/kg	-0.09047	0.1258	132	-0.72	0.4735
time*Diet	10,000 IU/kg	-0.04857	0.1174	132	-0.41	0.6799
time*Diet	100 IU/kg	0

Time Series Regression: Example

- How do we interpret the coefficients?
 - The *main effects* b_1 and b_2 represent the effect of the 1,000 and 10,000 IU/kg diets relative to the reference diet (100 IU/kg)
 - The *main effect* b_0 represented the tumor growth rate for the reference diet.
 - The *interaction terms* b_3 and b_4 represent the change in the growth rate of the 1,000 and 10,000 IU/kg diets relative to the reference diet (100 IU/kg)
 - Thus the growth rate for the 1,000 IU/kg diet is: $b_0 + b_3$
 - Thus the growth rate for the 10,000 IU/kg diet is: $b_0 + b_4$

Time Series Regression: Example

- In order to compare the growth rates between diets, we need to compare:
 - Diet 100 to Diet 1,000: b_0 versus $b_0 + b_3$
 - Diet 100 to Diet 10,000: b_0 versus $b_0 + b_4$
 - Diet 1,000 to Diet 10,000: $b_0 + b_3$ versus $b_0 + b_4$
- Programs like SAS (and perhaps SPSS) will allow us to make these specialized comparisons:

Contrast:	Test Statistic	p-value
Tumor Growth Rate: 1000 IU/kg vs. 100 IU/kg	0.52	0.473
Tumor Growth Rate: 1000 IU/kg vs. 10000 IU/kg	0.12	0.731
Tumor Growth Rate: 10000 IU/kg vs. 100 IU/kg	0.17	0.680

Survival Analysis

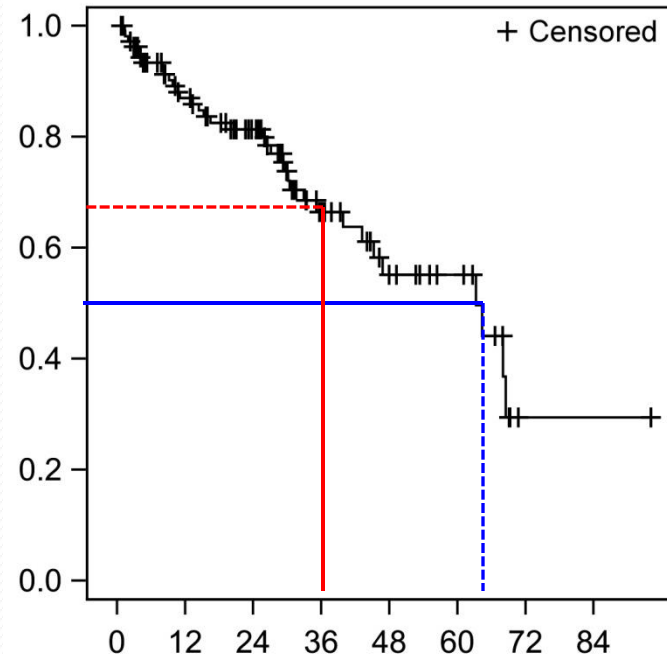
- General Purpose
 - In some studies you may be evaluating time-to-event outcomes.
 - In your study, not all patients:
 - Are observed for the same duration
 - Experience the event of interest
 - Ex: Survival and Recurrence data
 - How can we make comparisons between groups or evaluate relationships with respect to this type of data?
- Time to Event Data
 - Considered binary: (T, E)
 - T = time
 - E = event indicator
 - 1 = event occurred at time T
 - 0 = censored at time T (subject has been followed until time T with no event)

Survival Analysis

- Comparing Cohorts
 - Is there a difference in the distribution of a time-to-event outcome?
- **Kaplan Meier Methods**
 - Kaplan-Meier Curve summarizes the Survival Function
 - Survival Function: $S(t)$ = probability a subject makes it to time t without observing the event
 - Can obtain estimates of:
 - **Median Survival Time** - not appropriate to use the mean with censored data
 - **Survival Rates** – what proportion survived to a given time

Survival Analysis

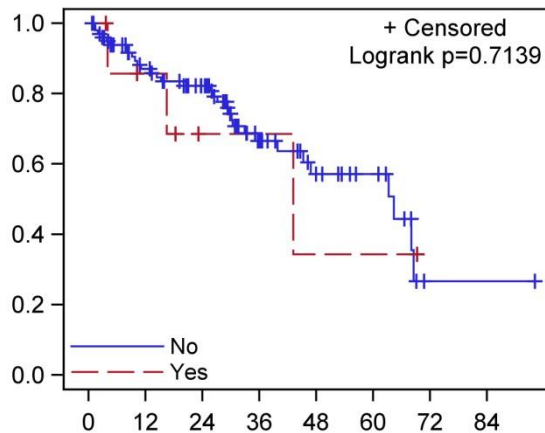
- Kaplan-Meier Curve
 - Median Survival
 - Where does the curve cross 50%
 - Survival Rates
 - What is the survival rate when you pass a given time point



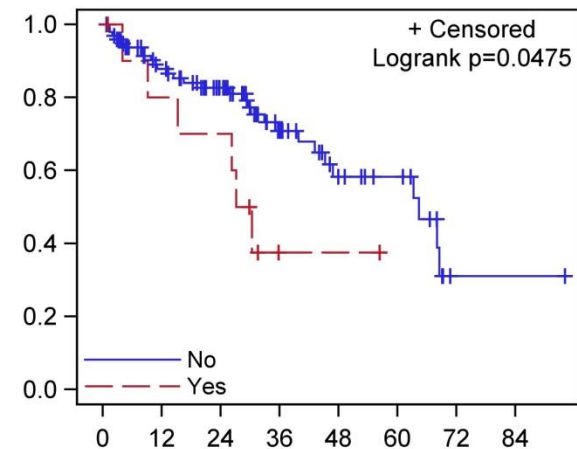
- If we have multiple cohorts, then we can produce a KP curve for each cohort
- The survival distributions are compared between cohorts using the **Log-rank Test**
 - Non-parametric test

Survival Analysis

- Kaplan-Meier Example
 - A study is conducted in order to examine clinical and surgical factors associated with survival in HIPEC treated patients.
 - Two of these factors were:
 - Neutropenia and Re-operation



Group	Censoring Fail/Cens/Total	Median Time (95% CL)	36 Month Survival Rate
No	32 / 72 / 104	64.4 (48.2, 88.5)	68.5% (53.8%, 78.9%)
Yes	3 / 5 / 8	43.2 (4.0, NR)	68.8% (21.3%, 91.2%)



Group	Censoring Fail/Cens/Total	Median Time (95% CL)	36 Month Survival Rate
No	29 / 72 / 101	64.4 (48.2, NR)	70.8% (57.9%, 80.4%)
Yes	6 / 5 / 11	28.8 (4.0, NR)	37.5% (9.6%, 65.0%)

Survival Analysis

- Regression for Time-to-Event data
 - Is there are relationship between a variable (measure) and the time to some event?
- **Cox Regression Models**
 - Models the survival as a function of a predictor(s)
 - Assumptions:
 - Proportional Hazards – the predictor effects the hazard rate in a proportional manner
 - Hazard Ratios:
 - The effect of a covariate on the hazard rate
 - *Relative Risk* of an event occurring in the next instance

Survival Analysis

- Cox Regression Example:
 - Consider the HIPEC study

Cohort	HR (95% CI)	P-value
No Re-Operation Required	1.00	-
Re-Operation Required	2.41 (1.01 – 5.90)	< 0.05

- Using Cox regression to model Survival as a function of Re-Operation Status (yes/no)

Statistics for Grants and Proposals

- For grants submissions, there are a few important statistical aspects that must be included for each Aim
 - Primary Objective and Primary Outcome
 - What is the main objective and outcome used?
 - “The primary objective is to evaluate the association between mTOR expression and immunosuppression”
 - What type of data?
 - “The primary outcome mTOR expression will be treated as ordinal data”

Statistics for Grants and Proposals

- Primary analysis
 - What methods will you apply?
 - Analysis plan
 - “The association between mTOR and immunosuppression will be assessed using a two-sided Wilcoxon signed rank test.”
 - Don't worry about listing what descriptive statistics you will use

Statistics for Grants and Proposals

- Power Justification
 - What effect size can you detect with 70 to 90% power?
 - “With a sample of 20 subjects per cohort, we have an 80% chance of detecting a 1.2 standard deviation difference between cohorts”
 - What power do you have to detect a given treatment difference or effect size?
- Useful Websites:
 - Power Calculations:
 - <http://www.sample-size.net/>
 - Statistical Analyses
 - <http://www.ats.ucla.edu/stat/dae/>

Statistics for Grants and Proposals

- Secondary Analysis
 - Are there any secondary objectives?
 - What is the analysis plan for those secondary objectives?
- Significance Level
 - What is your significance level?
 - Adjustments for multiple tests?
- Software
 - What software will you be primarily using?

Understanding Limitations

- Statistical
 - Not everything is as straightforward or simple as it seems
 - There are, unfortunately, a lot of nuances in statistics
- Software
 - Not all software can do the same things

Understanding Limitations

- Personal
 - Most analysis are straight forward and relatively simple
 - The more you do it, the more comfortable and easier it is
 - If you come across something you are unfamiliar with, remember two things:
 - Statistical software generally lets you run any analysis for your data, whether it is correct or incorrect
 - Its ok to ask for help

Biostatistics Core

- The Biostatistics Resource ensures that biostatistical, bioinformatics and biomathematical support is readily available to basic, clinical and population-oriented RPCI collaborators
- LIMS
 - <https://rpcilims.roswellpark.org/lims/logon.jsp>

Questions?