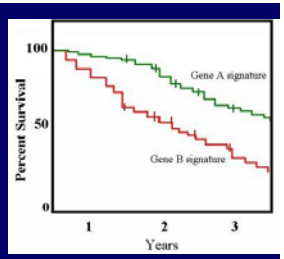
Clinical Curiosity

(...and how the Guinness Brewery changed history)

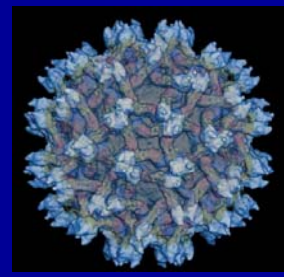
Katheryne Downes, M.P.H.
Statistical Data Analyst
Tampa General/USF College of Medicine

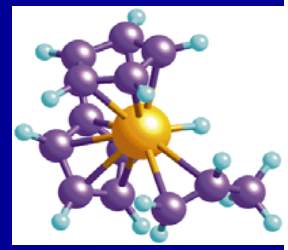
What is Biostatistics?



$$\hat{OR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{a_i d_i}{n_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)}$$

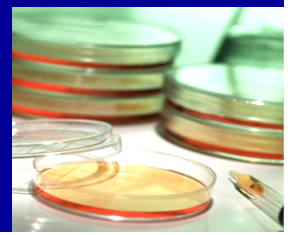
?





$$\hat{S}(t) = \prod_{t_i<t} \frac{n_i - d_i}{n_i}$$

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$



Who's done what?

- Literature Review
  - Don't want to duplicate efforts (or maybe you should?)
  - Can give ideas about how to conduct the study
  - Required for sample size calculations

Critical Review of Literature

- How were patients selected/recruited?
- What population are they attempting to generalize to?
- What was the sample size?
- Were sample size calculations completed prior to the study?
- What are the possible confounding variables? What was done to control for these variables?
- Definition of intervention?
- Definition of outcomes?
- **Statistics?**
- **Interpretation of findings and conclusions?**

Critical Review: Sample

- How were patients selected? (selection criterion)
- What population are the authors trying to generalize to?
- Sample sizes? (calculations?)

Critical Review: Treatments/Outcome

- What intervention/treatment is being investigated?
  - Dose
  - Devices
  - Length of administration
  - Randomization
- What outcomes are being investigated?
  - Definitions
  - Measures
  - Confounding variables
  - Conclusions justified by results?

"Notice all the computations, theoretical scribblings, and lab equipment, Norm. ... Yes, curiosity killed these cats."

Be Careful…

- REMEMBER: Just because it's published does not necessarily mean that it's a good study or that it's without flaw. Also- remember publication bias: Studies that show non-significant findings are often NOT published (Despite the fact that they are equally important)

Studies show that 34.7% of all statistics are made up on the spot….

Chicago Daily Tribune

HOME

DEWEY DEFEATS TRUMAN

G. O. P. Sweep Indicated in State; Boyle Leads in City

REPUBLICAN TICKET AHEAD OF 1944 VOTE

Town Is Telling Gives Trend

Tops Coghlan in Hot Race for Attorney

RECORD CITY VOTE SEEN IN LATE TALLIES

Suburban Ballot Near 75,000

BULLETINS ON ELECTIONS

Early Count Gives G.O.P. Senate Edge

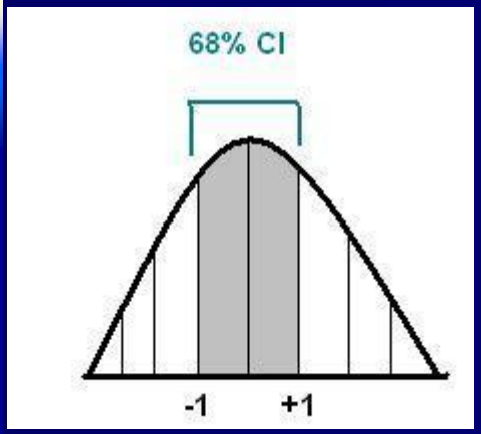
PUTS G.O.P. BACK IN THE WHITE HOUSE

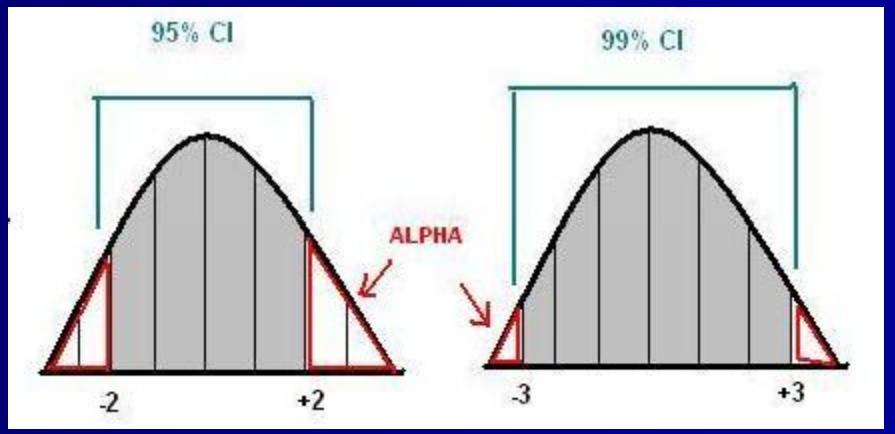
Stable Electoral Margin Seen

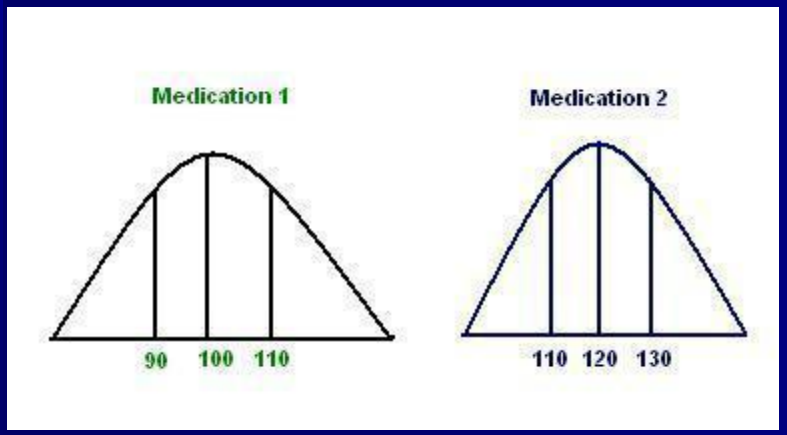
Statistics in Literature: The Basics

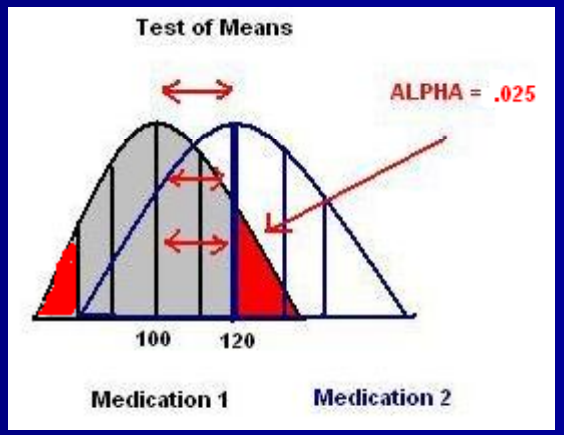
- Statistic
- Confidence Intervals
- Significance Values

Statistics in Literature: The Confidence Interval





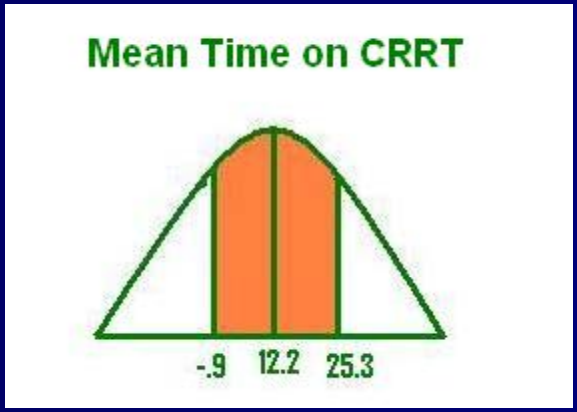


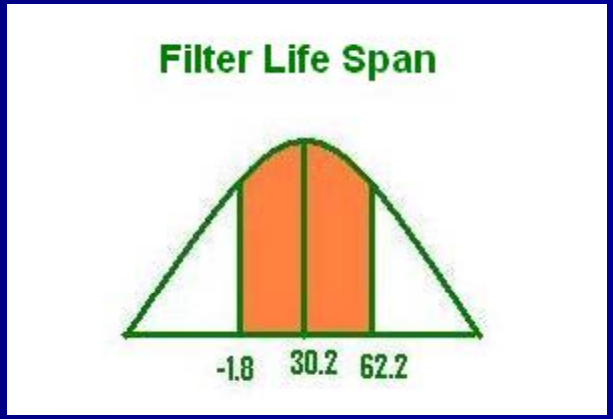


What's the Problem?

Filter Life Span during Citrate and Heparin Anticoagulation

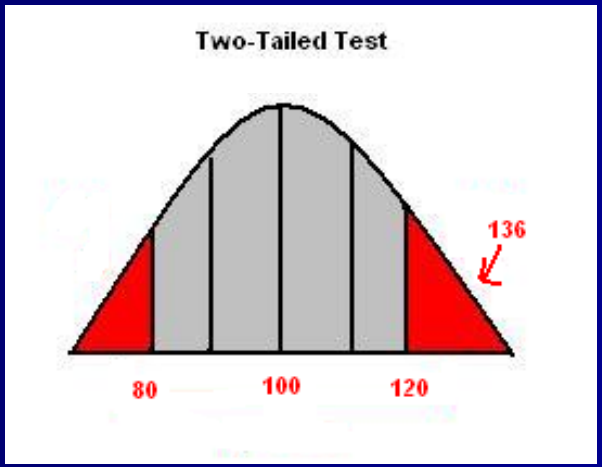
The mean time on CRRT in the overall group was 12.2 ± 13.1 days. Thirty-seven patients with ongoing bleeding or a high risk of bleeding were exclusively anticoagulated with citrate. Eighty-five patients received only heparin as anticoagulant. In the heparin-only group, 51 patients were treated by CVVH and 34 patients by CVVHD. The mean heparin dose during heparin anticoagulation was 530 ± 230 IU/h. Eighty-seven patients received low-dose heparin in combination with citrate for anticoagulation of the extracorporeal circuit. The mean heparin dose was 270 ± 130 IU/h ($p < 0.01$ vs. heparin). The aPTT sampled 24 h after the initiation of the CRRT was 50.1 ± 29 s in the heparin group, 48.1 ± 18 s in the low-dose heparin and citrate group and 52 ± 23 s in the citrate group (ns between the groups). The filter life span was significantly longer during citrate compared to heparin anticoagulation (80.2 ± 60 vs. 30.2 ± 32 h; $p < 0.001$). No difference was found between citrate and citrate-hep-





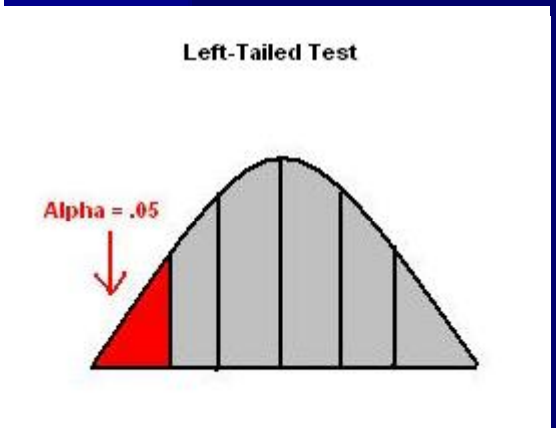
Statistics in Literature: The Confidence Interval

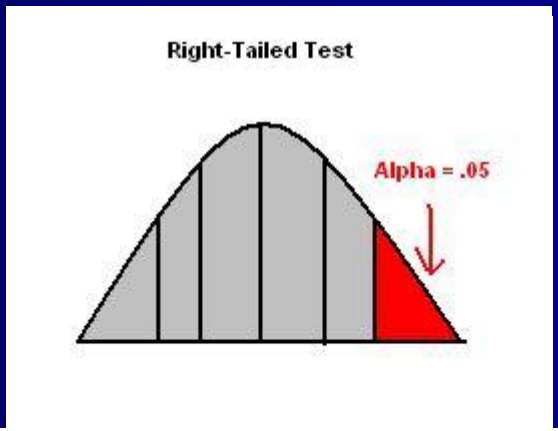
- Confidence Intervals
  - Estimation (Avg IQ = 100; 95% CI= 80-120)
  - Hypothesis Testing
    (Sample Avg IQ = 136, normal 95% CI = 80-120)

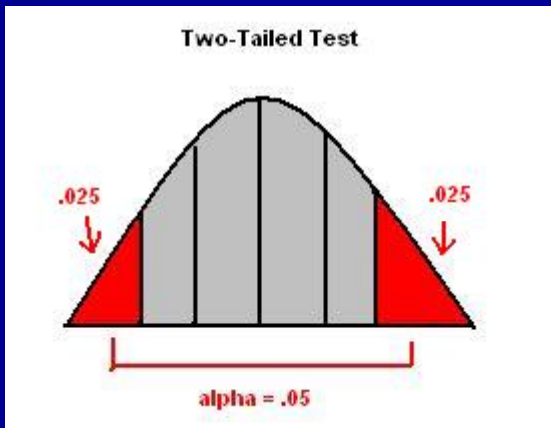


One-tailed or Two-tailed?

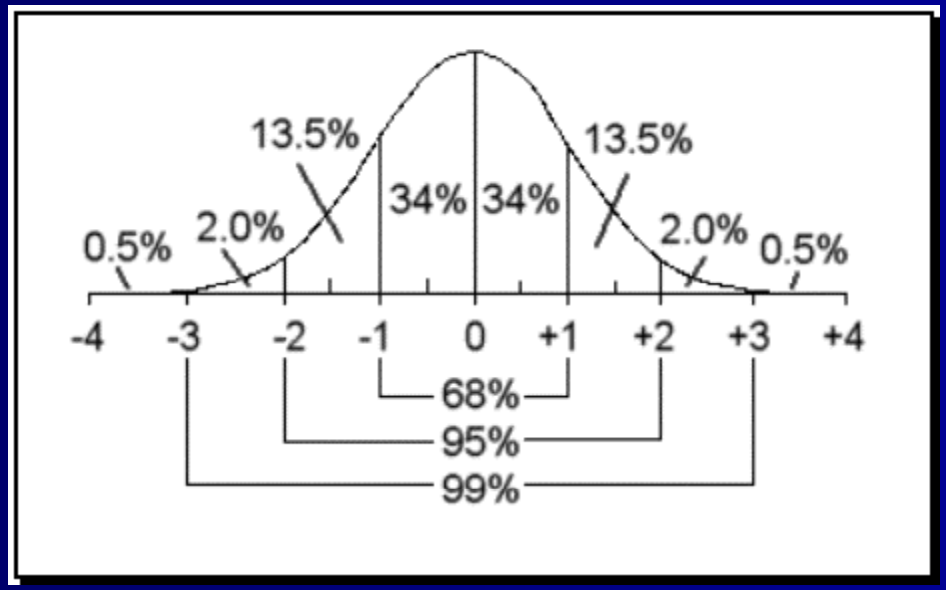
- One-tail:
  - We hypothesize Drug A is worse than Drug B
  - We Hypothesize Drug A is better than Drug B
- Two-Tailed:
  - We hypothesize Drug A performs differently than Drug B (direction isn't specified, more conservative test)







The Empirical Rule



Statistics in Literature: Significance Values

- P-value: the probability of observing your finding by chance alone.

  A p-value = .001 means that the probability of observing that particular event by chance would only be about 1/1000. Translation? You can be fairly certain that your observation did NOT occur by chance alone- something intervened.

Sample Size Calculations

- Depends on what test you're planning on conducting, but, in general…
  - Expected value in your control
  - Expected differences
  - Amount of variation known to exist

Sample Size Calculations: t-tests

From Literature/pilot study

- Standard deviation
- Expected difference (based off experience, previous research or other evidence)

Remember: select your numbers from a well-designed study. Be Careful!!

Sample Size Calculations: Proportions Test

From the literature/pilot study:

- Proportion of observed events in the control group
- Anticipated proportion of observed events in the active group (based off previous trends)

How the Guinness Brewery Changed History…

- "Student's" t-test
- William Gossett (left)
- R.A. Fisher (right)


'Student' in 1908

Understanding Statistics in Literature

- Are the statistics appropriate?
- What, exactly, does this really *mean?*
  - What does an odds ratio of 1.5 really mean?
  - Why am I looking for a "1" or a "0" in this confidence interval?
  - What does a significant ANOVA tell you? (for that matter, what's an ANOVA!?!?!)

T-test/Z-test

- What type of data? Group Means (continuous)
- Reported as? t-statistic/z-score & p-value
- What does it REALLY test?
  The difference in group distributions- in particular- the difference in group means.

T-test/Z-test Continued...

- T-tests are used when the sample size for each group is very small
- Z-tests utilize the normal distribution and can be used when the sample size is adequately large
- Not Appropriate for categorical data

ANOVA: Analysis of Variance

- What type of Data? Means (continuous)
- Reported as? F-Statistic, p-value
- What does it REALLY test?

  It compares the distributions of several groups simultaneously- it examines whether the amount of variation between groups is greater than that of within groups. A significant F-statistic tells you that the groups are not all equal, but it does NOT tell you which groups are different.

ANOVA

- Once a significant F-statistic is obtained, your next step would be to do individual tests comparing two groups at a time.
- Again, cannot be used for categorical data.

Chi-Square

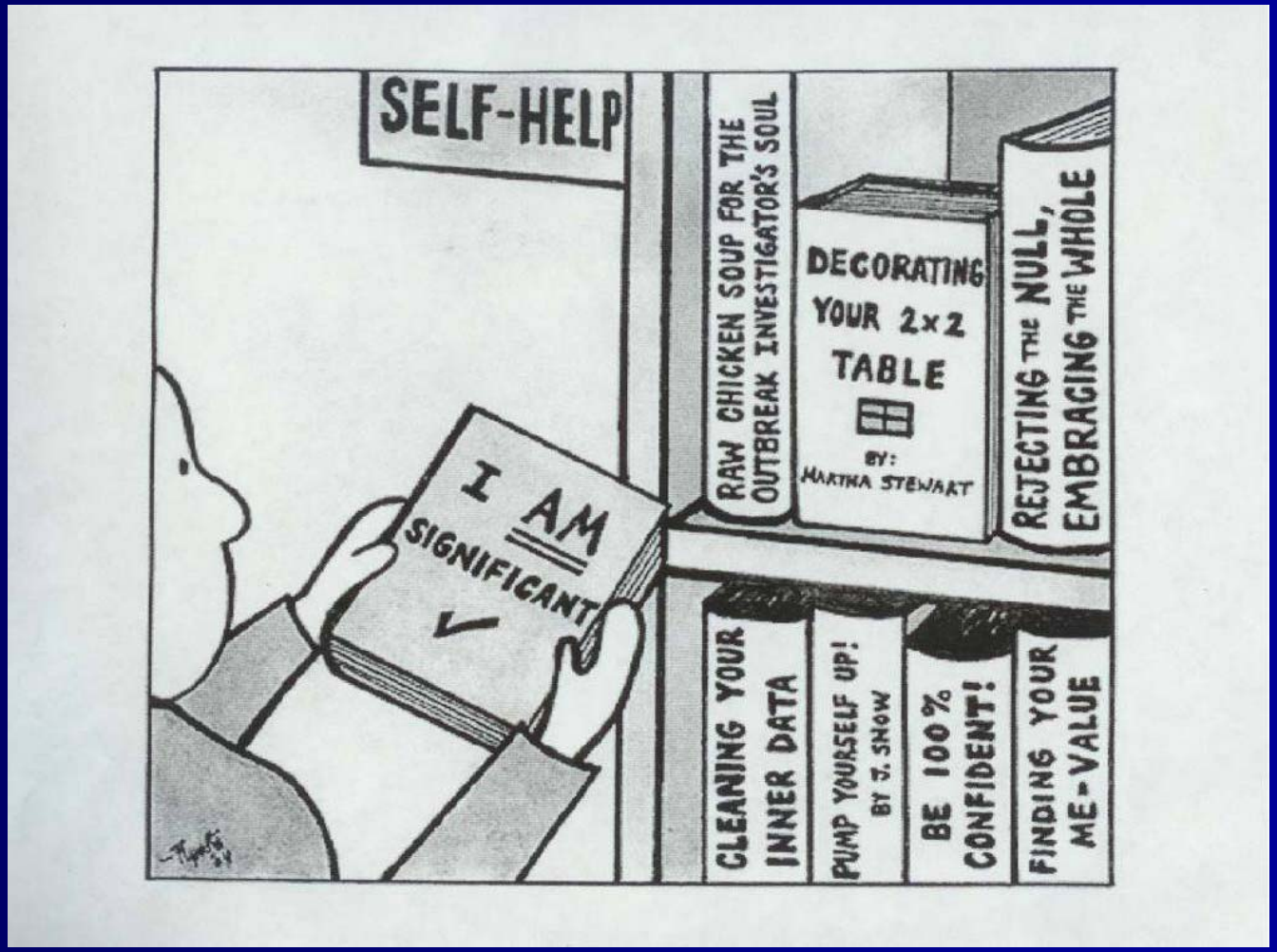
- What type of data? Categorical/dichotomous
- Reported as? X^2, p-value
- What does it REALLY test?

  A chi-square tests whether the observed frequency of an event is different than the expected frequency of the event (that which would occur by chance).
- ***Chi-Square tests can ONLY be used when each cell count is greater than or equal to "5"

Fisher Exact Test

- Works in basically the same manner as a chi-square, but it's used when you have cell counts below "5"
- An "exact" test CAN be used when cell counts are "5" or higher, but it becomes difficult to calculate with large sample sizes

Self-Help Books...



OR, RR, HR

- OR: Odds Ratio
- RR: Relative Risk or Risk Ratio
- HR: Hazard Ratio

- All three are ratios of risk- one test group is reflected in the numerator, the other in the denominator- therefore, if you get a ratio = "1" that means there's NO DIFFERENCE between groups. Keep this in mind while we look at them individually.

Relative Risk

- What type of data? Cohort Studies
- Reported as? RR, CI, p-value
- What does it REALLY test?
  The amount of risk associated with a particular exposure.
- ***Relative Risk can be safely used in cohort studies because we have incident rates available. However, the relative risk cannot be calculated in case-control studies.

RR: Interpretation

- RR* <1: Exposure is Protective
- RR*=1: No Difference
- RR*>1: Exposure is Risk Factor

- RR, CI, and p-value
  - RR = 1 = NO DIFFERENCE
  - What would a CI containing “1” mean?

(RR*: The same thing applies to OR and HR)

Odds Ratios

- What type of data? Case/Control Studies
- Reported as? OR, CI, p-value
- What does it REALLY test?
  Approximation of RR: The amount of risk associated with a particular exposure.
- ***An Odds Ratio must be used in case-control studies as the measure of risk because we have incomplete information about the prevalence/incidence of the disease in the calculations

Hazard Ratio

- What type of data? Survival
- Reported as? HR, CI, p-value
- What does it REALLY test?

  The ratio of two hazard rates-amount of "hazard" associated with being in a particular group. A hazard rate is the number of people dying divided by the number of people alive in each interval.

Hazard Ratio NOTES

- The Proportional-Hazard Assumption: The hazard rate between the two groups must remain proportional over time. If the hazard associated with one group changes with time- the hazard ratio CANNOT be used.

- HR = 2 ≠ 2x more likely to die, it means that at any given interval, twice as many people are dying in the intervention group compared to the control group. This might sound the same, but it's not- the first reflects probability, the second reflects rates in time. It's a subtle, but important difference.

- HR = 2.5768: 157.68% increase in the amount of hazard when compared to the other group.

Clinically Significant?

- Statistical Significance does NOT necessarily mean clinical significance!!
- EXAMPLE

Clinically Significant?!?

Results: A total of 218/711 (30.7%) of AAA patients undergoing EVAR had anemia at baseline. After adjustment for various risk factors, hemoglobin level was inversely related to maximum AAA diameter (β: − .144, 95%-CI: −1.482 − .322, $P = .002$). Post-EVAR survival was 65.5% at 5 years and 44.4% at 10 years. In long-term follow-up, survival was significantly lower in patients with anemia as compared to patients without anemia ($P < .0001$ by log-rank). Baseline hemoglobin levels were independently related to long-term mortality in multivariable Cox regression analysis adjusted for various risk factors (adjusted HR: 0.866, 95% CI: .783 to .958, $P = .005$). Within this model, statin use (adjusted HR: .517, 95% CI: .308 to .868, $P = .013$) was independently related to long-term survival, whereas baseline AAA diameter (adjusted HR: 1.022, 95% CI: 1.009 to 1.036, $P = .001$) was an independently associated with increased mortality.

Conclusions: Baseline hemoglobin concentration is independently associated with AAA size and reduced long-term survival following EVAR. Thus, the presence or absence of anemia offers a potential refinement of existing risk stratification instruments. (J Vasc Surg 2007;46:676-81.)

Linear Regression

- What type of data? Continuous Outcome
- Reported as? β, CI, p-value
- What does it REALLY test?

  Determines the amount of variation in the outcome that can be explained by the introduction of different explanatory variables.

Linear Regression

- Assumes that there is a LINEAR relationship between variables. If there isn't- Linear regression cannot be used. A simple X-Y plot can aid in preliminary determination of the nature of the data…
- Assumes that the outcome is continuous- if it's dichotomous, can't use linear regression.
- Again- beware of statistical significance vs. clinical significance.

Logistic Regression

- What type of data? dichotomous Outcome
- Reported as? β, CI, p-value
- What does it REALLY test?

  Determines the amount of variability in the outcome variable that can be explained by the introduction of different explanatory variables.


YES,
IN MY PROFESSIONAL OPINION, YOU'VE REGRESSED ENOUGH...

Regression Coefficient Interpretation

- "The regression coefficient (b_1) estimates the average increase in the dependent variables per unit increase in the independent variable…" (Szklo & Nieto)

- *WHAT?!?!*

Regression Interpretation

- Example:
- Y= number of items
- X= number of months
- Y = 5.13 +0.976x

- Breakdown:
- "5.13" : This really isn't meaningful, so you can just ignore it for our purposes…
- Positive slope indicates increase (positive relationship)
- It appears that for each additional month, the number of items is increasing by approximately "1" (.976)

Regression Interpretation

- Since beta reflects the slope of the line, $\beta = 0$, would reflect no relationship between the variable of interest and the outcome. Again, we can also check a CI for $\beta = 0$ as well.

- A regression equation cannot automatically be used to predict responses- the equation must be tested out in a new population to verify it's predictive capabilities.

- GIGO: Garbage IN, Garbage OUT

  - Calculate a regression line predicting the number of Starbucks coffees I need daily, using the number of ceiling tiles in this auditorium…

    …Yea, it could be done, but why would you do it?

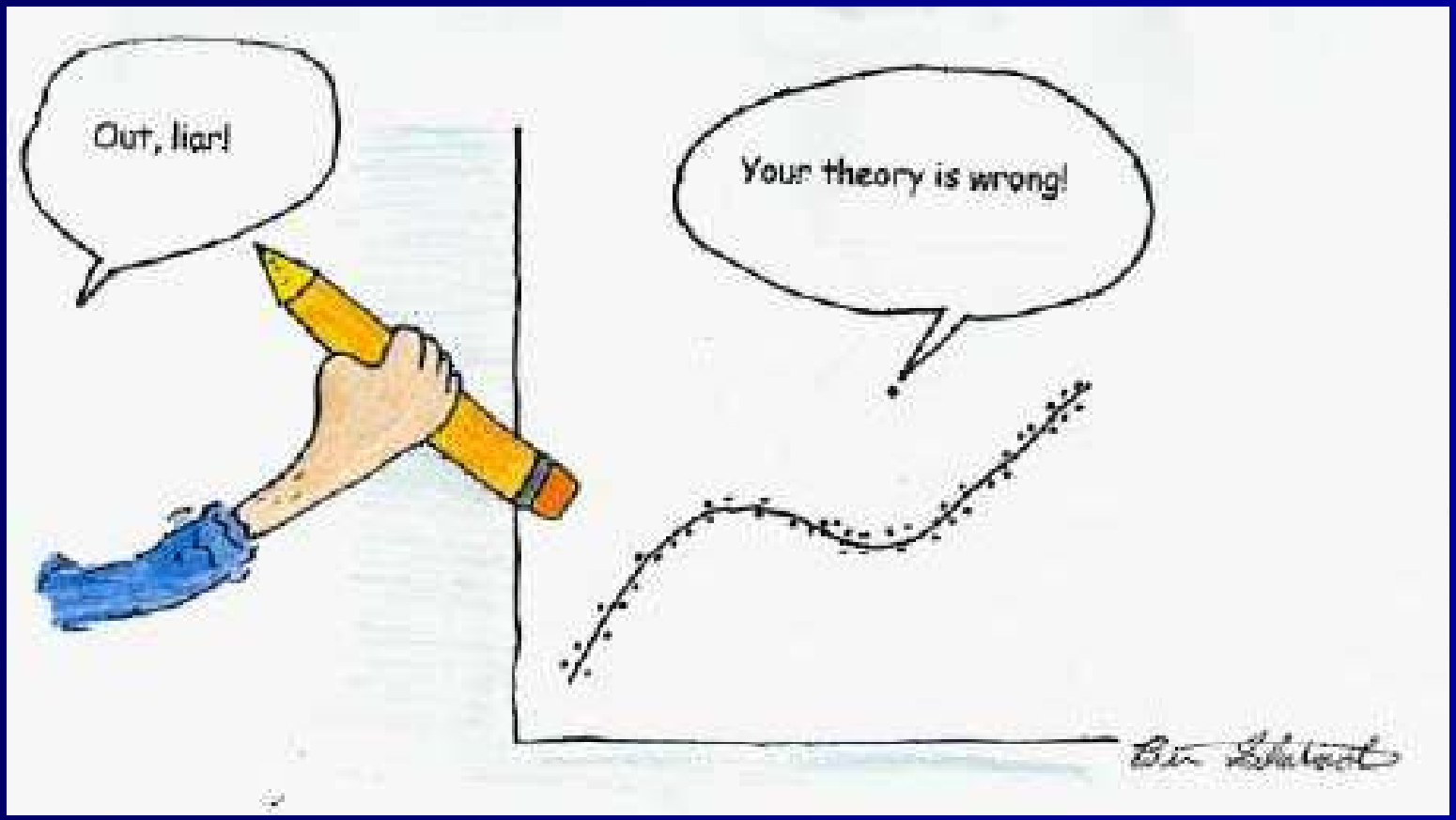
Survival Analysis

- Most survival analysis tests are done using adapted analytic methods that we've already covered: ANOVA, Chi-Square, etc.
- Be Cautious about how many people are in each group.
- Be aware that there may be unknown variables influencing your results (Beware the LURKING variables…)
- Censoring Issues
- Survival "Form"

REMEMBER!

- No matter what- if you find a significant result, there's still a small possibility that you're WRONG. This is inherent in probability- we don't have 100% certainty. We can only attempt to minimize the possible problems.

- If you fail to find a significant result- it doesn't necessarily mean that there isn't a relationship there. You might have structured the study wrong, used the wrong statistics, the wrong model, the relationship might not be the form that you think it is (linear regression on curvilinear data), or there might be another variable interfering that you don't know about…(damned lurking variables again…)


Out, liar!
Your theory is wrong!

On-Site Biostatistics: The Take-Home Menu

- Clinical Trial Design
- Database Design
- Sample Size Calculations
- Randomization Schemes
- Data Analysis
- Instruction
- IRB Statistical Review
- Publication consultation

Thank you!