F- Distribution

Skewed right

Mean is 1

Only non negative values

Gathering Data

Treatments: experimental conditions which correspond to assigned values of explanatory variable.

Observational studies: watch and observe values on response variable (non experimental)

Advantage of Experiments over Observational Studies: experiments reduce potential for lurking variable by random selection, also experiment is only way to determine causality

Sample Survey: selects sample from population and gathers data

Sampling Frame: list of subjects in the population from which the sample is taken

Simple random sampling: when each possible sample of that size has the same chance of being selected

To select a simple random sample: number the subjects in the sampling frame using numbers of the same length (number of digits). select numbers of that length from a table of random numbers or using a random number generator. include in the sample those subjects having numbers equal to the random numbers selected.

Margin of Error tells us how well the sample estimate predicts the population percentage. Ex. A survey results says margin of error is +/- 3% MEANS "it is very likely that the reported sample percentage is no more than #5 lower/higher than the population percentage

Bias: When certain outcomes will occur more often in the sample than they do in the population.

Sampling bias occurs from using nonrandom samples or having undercoverage.

Gathering Data (cont)

Nonresponse bias occurs when some sampled subjects cannot be reached or refuse to participate or fail to answer some questions.

Response bias occurs when the subject gives an incorrect response (perhaps lying) or the way the interviewer asks the questions (or wording of a question in print) is confusing or misleading.

Large Sample doesn't guarantee unbiased sample

Convenience Sample problems: results only apply to observed subjects, unlikely to be representative of population, often severe biases result

Key Parts of a Sample Survey:Identify the population of all subjects of interest. Construct a sampling frame which attempts to list all subjects in the population. Use a random sampling design to select n subjects from the sampling frame. Be cautious of sampling bias due to nonrandom samples (such as volunteer samples) and sample undercoverage, response bias from subjects not giving their true response or from poorly worded questions, and nonresponse bias from refusal of subjects to participate.

3 Components of Good Experiment:

1. Control Group - placebo, allows to analyze effectiveness

2. Randomization - eliminates researcher bias, balances comparison groups on known and lurking variables

3. Replication: allows to attribute observed effects to tx rather than regular variability

Statistically Significant: if observed difference is larger than would be expected by chance

Can generalize only to population represented by sample

14.1 One-Way Anova: Comparing Several Means

One way ANOVA is an ANOVA with a single factor

Factor: categorial explanatory variable

Test analyzes whether differences observed among the *sample* means could have reasonably occurred by chance, if the null hypothesis of equal *population* means were true

Evidence against the null is stronger when the variability between sample means increases and as the sample sizes increase

Assumptions and the effects of violating them:

Population distributions are normal (Moderate violations of the normality assumption are not serious.) These distributions have the same standard deviation. (Moderate violations are not serious.) The data resulted from randomization.

Misleading results may occur with the F-test if the distributions are highly skewed and the sample size N is small.

Misleading results may also occur with the F-test if there are relatively large differences among the standard deviations (the largest sample standard deviation being more than double the smallest one).

Several T-Tests vs. F-test: If separate t tests are used, the significance level applies to each individual comparison, not the overall type I error rate for all the comparisons. However, the F test does not tell us which groups differ or how different they are.

One Way ANOVA example

Question: Three groups, with different French skills, scored on one quiz

Assumptions

> Independent Random Samples, normal population distributions with equal standard deviations

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
http://crosswordcheats.com

One Way ANOVA example (cont)

Hypotheses

> H0: u1=u2=u3 Ha: at least two population means are unequal

Test statistic

> F= btwn groups variability/within groups variability
> df1 = (g-1) df2 = (N-g)

P-Value

> Right tail probability of above observed F value

Conclusion

> Interpret in context, reject Ho based on p-value being below or = significant value

14.2

Confidence Intervals Comparing Pairs of Means

s is square root of within groups variance estimate (s^2)

For 95% Confidence Interval comparing means ui - uj: when the confidence interval does NOT containt 0, we can infer the population means are different, the interval shows just how different they may be

Example: for comparing the very happy and pretty happy categories, the confidence interval for u1 - u2 = (0.7, 5.3)
Since the CI contains only positive numbers, this suggests that on average people who are very happy have more friends than people who are pretty happy

Effects of violating assumptions
When the sample sizes are large and the ratio of the largest standard deviation to the smallest is less than 2, these procedures are robust to violations of these assumptions.
If the ratio of the largest standard deviation to the smallest exceeds 2, use the confidence interval formulas that use separate standard deviations for the groups.

Tukey Multiple Comparison

14.2 (cont)

Ex. Groups: (Very Happy, Pretty Happy) Difference of Means: (u1-u2) 95% CI: (0.7, 5.3) *Tukey 95% Multiple Comparison* (0.3, 5.7)

The Tukey intervals hold with an *overall* confidence level of 95%, this confidence applies to all intervals. Tukey is wider than separate CI's because uses a higher confidence level to achieve 95% for all intervals.

The Tukey Confidence interval for u1-u2 contains only positive values so infer that u1>u2, mean number of good friends higher for very happy than pretty happy (but maybe barely so).

ANOVA and Regression

Two -Way ANOVA

Two -Way ANOVA (cont)

If the evidence of interaction is not strong (that is, if the P-value is not small), then test the main effects hypotheses and/or construct confidence intervals for those effects.

Repeated Measures ANOVA

Sum of Squares in One Way Repeated Measures

Independent Groups:	Dependent Groups
SS Groups (df = g-1)	SS Groups (df = g-1)
SS Error (df = N - g)	SS subjects (df = subj - 1)
	SS error (df = n-g-subj.+1)

In repeated measures (dependent groups) ANOVA, the variability of the subjects is calculated (as if it was a factor) and is not included in the error sums of squares.

A very important assumption underlying repeated measures ANOVA is sphericity and, relatedly, compound symmetry. When either of these assumptions are violated, the P-values tend to be too small. A Greenhouse-Geisser adjustment to the dfs will accommodate for any potential violations of this assumption.

Two-factor studies often have different (i.e., independent) samples on one of the factors and the same (i.e., dependent) samples on the other factor. The factor with different groups of subjects is called the “between--subjects” factor and the factor with repeated measures is called the “within-subjects” factor.

Difference Between 1 and 2 way ANOVA

One way analyzes relationship between mean of quantitative response variable and groups that are categories of a factor

Two Way ANOVA analyzes quantitative response variable on two categorical response variables

Null Hypothesis In two-way ANOVA, a null hypothesis states that the population means are the same in each category of one factor, at each fixed level of the other factor.

Ex. Ho: Mean corn yield is equal for plots at the low and high levels of manure, for each fixed level of fertilizer. From the output, you can obtain the F-test statistic of 6.88 with its corresponding P-value of 0.018. The small P-value indicates strong evidence that the mean corn yield depends on manure level.

No interaction between two factors means that the effect of either factor on the response variable is the same at each category of the other factor.

Usually test hypothesis that there is no interaction first

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
http://crosswordcheats.com