Definitions:

Study (experimental) unit (subject) - one member of a set of entities being studied.
Parameter (also, estimand) -proportion of childless households in the population.
Estimate - proportion of childless households in the sample.
Parameter is a population characteristic.
Estimand is a parameter that is being estimated (**μ**).
Statistic is a function of the data (therefore, it is random).
Statistic: A function of the data, **μ**= **μ** (y). Test Statistic: Statistic used to weigh evidence supporting and contradicting the null hypothesis.
Reference Distribution: Probability distribution of the test statistic, assuming that the null hypothesis istrue, f(**μ** (Y)|H0).
Estimator is a statistic used as a guess for the value of the estimand ( ).
Estimate is a quantity, which is a particular realization of the estimator (4/12 = 0.33).
Target Population: A collection of units a researcher is interested in; a group about which the researcher wishes to draw conclusions.
Sampling frame: Collection of units that are potential members of the sample.
Overcoverage: selecting units that are outside your sampling frame.
Undercoverage: excluding units that are within your sampling frame.
Sample: A [randomly selected] subset of a sampling frame
Simple Random Sampling (SRS) – every subset of n units has equal chance to be selected
Stratified Sampling – split the population into homogeneous subpopulations and use SRS (or another method) within a sampling frame of each subpopulation.
Systematic Random Sampling - select every k'th unit from the ordered sampling frame, starting randomly from of the first k positions.
Variable probability sampling – allow units to have nonequal probabilities of being sampled.
p-value - probability that the test statistic would be at least as extreme as observed, under the null hypothesis. (or "strength of evidence against the H0")
Significance level (α) is the criterion compared against the p-value. The null hypothesis is rejected if p-value is lower than α.

α reflects the probability of rejecting the null hypothesis given that it is true (**Type I error**).

Scope of inference

Internal validity: are assumptions of the test satisfied?
External validity: possible to make inference to a broader population?

Sampling distribution of a statistic is a (reference) distribution that arises from a chance mechanism used to select a random sample from a population.
Type I Error (or significance level α): Probability of rejecting the null hypothesis, when the null hypothesis is true.
Type II Error (β): Probability of failing to reject the null hypothesis, when the null hypothesis is false.
Power (1-β): Probability of rejecting the null hypothesis, when a particular alternative hypothesis is true.
If alpha = 0, beta = 1; if alpha = 1, beta = 0; if alpha increases, beta decreases (power increases).
Increase power by: increase alpha; do one-tailed test; increase effect size; increase sample size
Most to least power, if assumption of t-tools is met

For small samples:
1. T-tools
2. Permutation test
3. Rank-sum test (or signed-rank test for paired data)
4. (Sign test for paired data)

Large samples rejection rates are basically equivalent.

No

Population Variance: $\sigma^2 = E\left((Y-\mu)^2\right)$
Sample Variance:

Sample variance is chi-sq distributed: $$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$$

Let Z0~N(0,1), independent of variance X^2~X^2_n, then $\frac{Z_0}{\sqrt{X^2/n}} \sim t_n$

Chi-sq distribution with n df: $$X^2 = \sum_{i=1}^{n} Z_i^{\,2} \sim \chi_n^2$$

Pooled sample variance: $$S_p^2 = \frac{(n_X-1)S_X^2 + (n_Y-1)S_Y^2}{n_X+n_Y-2}$$, where

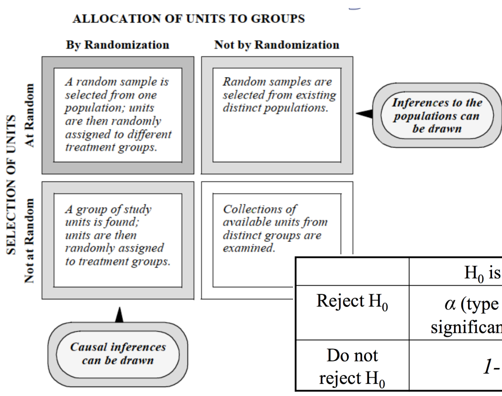
$$S_X^2 = \frac{\sum_{i=1}^{n_X}\left(X_i - \bar{X}\right)^2}{n_X-1},\quad S_Y^2 = \frac{\sum_{i=1}^{n_Y}\left(Y_i - \bar{Y}\right)^2}{n_Y-1}$$

Sampling distribution of pooled sample variance: $$S_p^2 \sim \frac{\sigma^2}{n_X+n_Y-2}\chi^2_{n_X+n_Y-2}$$

F-distribution: if $X \sim \chi^2_{n_X}$ and $Y \sim \chi^2_{n_Y}$, independent of each other, then

$$\frac{X/n_X}{Y/n_Y} = R \sim F_{n_X, n_Y}$$

(F-distribution with nx and ny d.f)



	H_0 is true	H_0 is false
Reject H_0	α (type I error, significance level)	$1\text{-}\beta$ (power)
Do not reject H_0	$1\text{-}\ \alpha$	β (type II error)

Sum of Squared Residuals	General Name	Calculation	Degrees of freedom	Distribu-tion	Corresponding variance (also called "Mean Square")
SSR_{full} Residual sum of squares (full)	SSW (Within Groups)	$\sum_{i=1}^{I}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_i\right)^2$	n-I	$\sigma^2\chi^2_{n-I}$	S_p^2 MSR_{full}
ESS (Extra sum of squares)	SSB (Between Groups)	$\sum_{i=1}^{I} n_i\left(\bar{Y}_i - \bar{Y}\right)^2 = \boxed{\sum_{i=1}^{I} n_i\left(\bar{Y}_i\right)^2 - n\left(\bar{Y}\right)^2}$	I-1	$\sigma^2\chi^2_{I-1}$	MSB
$SSR_{reduced}$ Residual sum of squares (reduced)	SST (Total)	$\sum_{i=1}^{I}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}\right)^2$	n-1	$\sigma^2\chi^2_{n-1}$	S^2 Variance of the entire sample as one group

(Fisher's) Randomization Test: is a distribution-free test for treatment effect in **randomized experiments.**
H0 : Zero treatment effect for all units, δ=0. Each unit's outcome is the same, regardless of the treatment assigned. Consequently, the distribution of outcomes is identical in two groups
Ha : Non-zero treatment effect for ALL units, δ≠0 .
Assumptions:

1. Random assignment to groups.
2. Under the H0, independence of study units.

Test statistic: Difference between average outcomes in the two groups*. $\bar{Y}_c - \bar{Y}_v$
*In fact, any summary of the data can represent a test statistic.
Randomization distribution is a reference distribution of a test statistics in a randomization test, where variation is due to random assignment of the treatment.

Permutation test is a distribution-free nonparametric test for association between group status and outcome in **observational studies.**
H0 : Outcomes not related to group status.
Ha : Outcomes are related to (associated with) group status.
Assumptions: Independence of study units.
Test statistic: Difference in average observed outcomes
between the two groups (or, any other statistic),
Permutation distribution is a reference distribution of a test statistics in a permutation test.
NOTE: Permutation tests are actually evaluating whether the distributions are equivalent, not just the

How does CLT apply?

 # checking assumptions:
Independence: plot subgroups, and see if relationship still holds (cluster effect)
Plot data over time (serial effects)
Plot data vs space (spatial effect)
Normality: Histograms should look normal – overlay with normal curve and kernel density plot
Boxplot or Q-Q plot will show outliers
Equal Variance: divide variances to see if ratio is <0.5 or > 2

T-test is robust:
-if populations are symmetrically skewed
-if sample sizes are equal (and large)
-when sample sizes are equal, pooled-ttest is robust to unequal variances
Sensitive: When the sample sizes are not equal, t-tests are more sensitive to skewedness and long-tailedness.
For small samples, t-tests are somewhat sensitive to markedly different skewedness in two groups.
Watch out for outliers.

Robustness of F-tests:
Not resistant to outliers. Fairly robust to non-normality. Independence w/I and b/w groups is necessary. Equal variance is necessary.

F-Test Assumptions:
Equal variances is crucial
Not resistant to outliers
Normality is not critical.

Confidence intervals: Yi ~iid N(mu, sigma^2)

One Sample t-test: $$\left(\bar{Y} - t_{1-\alpha/2,n-1}\frac{S}{\sqrt{n}},\ \ \bar{Y} + t_{1-\alpha/2,n-1}\frac{S}{\sqrt{n}}\right)$$

Wilcoxon Signed-Rank Test for Paired Data:
(more assumptions, but higher power than sign test)
Hypotheses: General formulation:
H0: The rank of the magnitude of within-pair difference is unrelated to the sign of the difference.
Ha: The rank of the magnitude of within-pair difference is related to the sign of the difference.

If differences are continuous r.v.'s:
H0: Median(Di)=0
Ha: Median(Di)≠0 (or >0, or <0)

If the distribution of differences is symmetric :
H0: E(Di)=0
Ha: E(Di)≠0 (or >0, or <0)

1. Calculate differences (+ or -).
2. Rank the absolute values of differences |Di|
3. $Z_i = \mathrm{Rank}(|D_i|),\ i = 1,\ldots,n.$

Test Statistic: (sum of ranks for positive differences)

$$S = \sum_{i=1}^{n} Z_{i|D_i>0}$$

Exact sampling distribution: do permutation test, and switch group status for each pair.
Normal approximation (for large sample sizes, m is >= 20, where m is the number of non-zero differences):

$$Z = \frac{S - m(m+1)/4}{\sqrt{m(m+1)(2m+1)/24}} \doteq N(0,1)$$

Assumptions:

1. Data are paired and come from the same

equality of the means:

One Sample z-Test
H0: **μ = μ0**
HA: **μ != μ0**

Test Statistic: $$Z = \frac{\bar{Y} - \mu_0}{(\sigma / \sqrt{n})}$$ (make sure to double p-value if doing 2-sided test, of ½ alpha)
Exact sampling distribution of Z under H0: Z~N(0,1)
Assumptions:
1. Independence
2. Known population variance, σ^2
3. Normality (CLT allows for deviations!)
4. Random sampling from population

One Sample t-Test
H0: **μ = μ0**
HA: **μ != μ0**

Test Statistic: $$T = \frac{\bar{Y} - \mu_0}{(S / \sqrt{n})}$$
Exact sampling distribution of T under H0: T ~t_{n-1} (t-distribution with df = n-1)
Assumptions:
1. Independence
2. Normality (CLT allows for deviations!)
3. Random sampling from population

Two Sample z-Test
H0: **μx – μy = μ0**
HA: μx – μy != μ0

Test Statistic: $$Z = \frac{(\bar{X} - \bar{Y} - \mu_0)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} \qquad Z \sim N(0,1)$$

Pooled Two-Sample t-Test
H0: μx – μy = μ0
HA: μx – μy != μ0

Test Statistic: $$T = \frac{(\bar{X} - \bar{Y} - \mu_0)}{S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}$$
Exact sampling distribution of T under H0: T ~$t_{nx+ny-2}$ (t-distribution with df = nx + ny -2)

Note:s $(\bar{X} - \bar{Y}) \sim N\left(\mu_X - \mu_Y, \sigma^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)\right)$,

$\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}$ is an unbiased estimator of $\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}$

Assumptions: same as Welch's, but
If using pooled variance – assume equal variances between populations.

Unpooled (Welch) Two-Sample t-Test
H0: μx – μy = μ0
HA: μx – μy != μ0

Test Statistic: $$T = \frac{(\bar{X} - \bar{Y} - \mu_0)}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}$$
Approximate Sampling distribution:
$$T \sim t_{\tilde{n}}, \text{ where } \tilde{n} = \left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2 / \left(\frac{s_X^4}{n_X^2(n_X - 1)} + \frac{s_Y^4}{n_Y^2(n_Y - 1)}\right)$$
Assumptions:
1. Independence between units:
   within each population (Yi with Yj and Xi with Xj)
   between populations (Yi with Xj)
2. Homogeneity of units within each population:
   Equal means within each population;
   Equal variances within each population;
   If using pooled variance – assume equal variances between populations.
3. Population distributions are Normal:
   CLT allows for deviations.
   3. Random sampling from populations.

Paired t-Test
$$X_i - Y_i = D_i \overset{i.i.d.}{\sim} N(\mu_D, \sigma^2)$$
H0: μD = μ0
HA: μD != μ0

Test Statistic: $$T = \frac{\bar{D} - \mu_0}{(S_D / \sqrt{n})}$$
Exact sampling distribution of T under H0: T~t_{n-1} (where n is # of pairs)
Assumptions same as one-sample t-test

F-test for equal variance (could also use levene's test): (caution: will reject for large samples when ratio is close to 1)

Two Sample t-test:

- Two-sample *t*-tests

Standard Error

$$\left(\bar{X} - \bar{Y} - t_{1-\alpha/2, df} \cdot SE,\ \ \bar{X} - \bar{Y} + t_{1-\alpha/2,\ df} \cdot SE\right)$$

- Pooled *t*-test: $SE = S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}},\ df = n_X + n_Y - 2$

- Unpooled *t*-test: $SE = \sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}},\ df = \tilde{n}$
$$\tilde{n} = \left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2 / \left(\frac{s_X^4}{n_X^2(n_X - 1)} + \frac{s_Y^4}{n_Y^2(n_Y - 1)}\right)$$

Two-Sample t-test for I samples (I > 2)

- Random sampling.

Pooled using *I* samples!

Confidence intervals: $SE = S_p \sqrt{\frac{1}{n_m} + \frac{1}{n_k}},\ df = n_1 + n_2 + \ldots + n_I - I$

$$\left(\bar{Y}_m - \bar{Y}_k - t_{1-\alpha/2, df} \cdot SE,\ \ \bar{Y}_m - \bar{Y}_k + t_{1-\alpha/2,\ df} \cdot SE\right)$$

Transformations:
Log Transformation:
1. **Randomized Experiments** (T=Treatment, C = Control)
$$Z_{T,i} = \ln(Y_{T,i}),\ i = 1,\ldots,n_T$$
$$Z_{C,i} = \ln(Y_{C,i}),\ i = 1,\ldots,n_C$$
Responses:
$$Z_{T,i} = Z_{C,i} + \delta \Leftrightarrow \frac{Y_{T,i}}{Y_{C,i}} = e^{\delta}$$
$$\exp\left(\bar{Z}_T - \bar{Z}_C\right) \text{ estimates } \frac{Y_{T,i}}{Y_{C,i}}$$
Interpretation: -- The response of an experimental unit to the treatment will be exp(ZbarT – ZbarC) times as large as its response to the control
2. Observational studies with **random sampling**:
$$\exp\left(\bar{Z}_2 - \bar{Z}_1\right) \text{ estimates } \frac{m_2}{m_1}$$
Interpretation: -- The median of the second population is exp(Zbar2 – Zbar1) times as large as the median of the first population.
3. In paired t-test: $Z_i = \ln(Y_{2,i}) - \ln(Y_{1,i}),\ i = 1,\ldots,n$
Randomized Experiments:
$\exp(\bar{Z})$ estimates a multiplicative treatment effect $\frac{Y_{2,i}}{Y_{1,i}}$
Observation studies w/ rand. Sampling:
$$\text{Median}(Y_{2,i} / Y_{1,i}) = m, \text{ then } \exp(\bar{Z}) \text{ estimates } m.$$

Square Root Transformation sqrt(Y): Good for moderately skewed data, but hard to interpret results
Reciprocal Transformation 1/Y: good for severely skewed data, can be used with negative data
Logit log(Y/(1-Y)) or arcsin(2Y-1): good for proportions

Nonparametric alternatives to t-tools:

Rank Sum test:
Hypotheses:
H0: (if treatments are randomized): Same as fisher's randomization test
H0: (if treatments are not randomized): same as permutation test
* If shapes and spreads of the two populations are similar: HA: There is a difference in medians between the two populations.
Transform data:
$$Z_{1,i} = \text{Rank}(X_i \mid X_1,\ldots,X_{n_X},Y_1,\ldots,Y_{n_Y}),\ i = 1,\ldots,n_X$$
$$Z_{2,i} = \text{Rank}(Y_i \mid X_1,\ldots,X_{n_X},Y_1,\ldots,Y_{n_Y}),\ i = 1,\ldots,n_Y$$
Ties are averaged

Test statistic: $$T = \sum_{i=1}^{n_X} Z_{1,i}, \text{ assuming } n_X \leq n_Y$$
Assumptions: Independence

Normal approximation to rank Sum test (if no ties, and samples greater than 10):
$$T \sim N\left(\frac{n_X(n_X + n_Y + 1)}{2}, \frac{n_X n_Y(n_X + n_Y + 1)}{12}\right)$$
With ties, Rbar is sample mean (mean of all ranks) and Sr^2 is sample variance for combined set of nx + ny ranks:
$$T \sim N\left(n_X \bar{R}, S_R^2 \frac{n_X n_Y}{n_X + n_Y}\right)$$

We can use this Z-statistic: $$Z = \frac{T - n_X \bar{R}}{S_R \sqrt{\frac{n_X n_Y}{n_X + n_Y}}}$$

Sign test for paired data (lower power than Wilcoxon signed-rank test)
H_0: $P(D_i>0)=P(Y_i>X_i)=0.5$,
General formuation: H_a: $P(D_i>0) \neq 0.5$ (or >0.5, or <0.5).

population.
2. Each pair is chosen randomly and independently.
3. The data are measured at least on an ordinal scale, but need not be normal.

Two-Sample t-Test for I samples
H0: μm – μk = μ0
HA: μm – μk != μ0

Test Statistic:
$$T = \frac{(\bar{Y}_m - \bar{Y}_k - \mu_0)}{S_p \sqrt{\frac{1}{n_m} + \frac{1}{n_k}}}; \qquad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \ldots + (n_I - 1)S_I^2}{n_1 + n_2 + \ldots + n_I - I}$$

Exact sampling distribution of T under H0 is:
T~$t_{n1+n2+\ldots+nI-I}$, where I is number of groups.
Assumptions:
Independence of units within each group;
Independence of units between groups;
Normality of all I populations;
Equal variances for all I populations;
Random sampling.

Anova F-test for Equality of All means in I samples
H0: μ1 = μ2 = …=μi (reduced model – equal means model)
Ha: at least one mean is different from the others (full model – separate means model)

SSR$_{full}$ = SSW (within groups)
$$S_p^2 = \frac{\text{SSR}_{\text{full}}}{n - I}$$
$$S_p^2 \sim \frac{\sigma^2}{n - I}\chi_{n-I}^2 \Rightarrow \text{SSR}_{\text{full}} \sim \sigma^2 \chi_{n-I}^2, \text{ d.f.} = n\text{-}I$$

SSR$_{reduced}$ = SST (**total error**)
S^2 is the sample variance of the entire sample, taken as one group:
$$S^2 = \frac{\text{SSR}_{\text{reduced}}}{n - 1}$$
$$S^2 \sim \frac{\sigma^2}{n - 1}\chi_{n-1}^2 \Rightarrow \text{SSR}_{\text{reduced}} \sim \sigma^2 \chi_{n-1}^2, \text{ d.f.} = n\text{-}1$$
ESS = SSB (between groups)
ESS = SSR$_{reduced}$ - SSR$_{full}$
$$\text{ESS} \overset{H_0}{\sim} \sigma^2 \chi_{I-1}^2, \text{ d.f.} = I\text{-}1$$
Test Statistic:
$$R = \frac{\text{ESS} / (I - 1)}{\text{SSR}_{\text{full}} / (n - I)} = \frac{\text{MSB}}{\text{MSR}_{\text{full}}}$$
MSB = Mean square between
MSR$_{full}$=Mean square residuals (within groups)
Easy way to calculate: MSR$_{full}$ == S_p^2
Exact sampling distribution of R under H0:
$$F_{(I-1, n-I)}$$

Assumptions (check w/ Q-Q or residual plot):
1. Normality of populations;
2. Equal population variances for all groups;
3. Independence within each group;
4. Independence between groups;
5. Homogeneity within each group;
6. Random sampling

Kruskal-Wallis Test (Nonparametric ANOVA)
H0: Median(Y1) = Median(Y2) = …=Median(YI)
HA: at least one median is different from the others

Test Statistic – convert to ranks, and do ANOVA on ranks:
$$KW = \frac{\text{ESS (on ranks)}}{S_R^2} \overset{H_0}{\sim} \chi_{I-1}^2$$
S_R^2 is the sample variance of **all** ranks (not pooled)
Assumptions:
1. Independence.
2. Random Sampling

F-test for Equality in a subset of groups

H_0: $\sigma_X^2 = \sigma_Y^2$ H_A: $\sigma_X^2 \neq \sigma_Y^2 \Leftrightarrow \sigma_X^2 / \sigma_Y^2 \neq 1$ Test Statistic: $F = \frac{S_X^2}{S_Y^2} \sim F_{n_X-1, n_Y-1}$ Same assumptions as t-test: In testing the equality of population variances, two assumptions are required: independent samples and normally distributed populations. 2-sided p-value is just 2*the one-sided value.	**If differences are continuous rv's:** H_0: Median(D_i)=0, H_a: Median(D_i)≠0 (or >0, or <0). **If differences are symmetric:** H_0: E(D_i)=0, H_a: E(D_i)≠0 (or >0, or <0) **Test Statistic: K = number of positive differences.** **Exact distribution of K: K ~Binom(m, 0.5), where m is final number of pairs with nonzero Di.** **Normal approximation (for large samples):** $Z = \frac{K - m/2}{\sqrt{m/4}} \dot\sim N(0,1)$ **Assumptions:** 1. The differences D_i are assumed to be independent. 2. Each D_i comes from the same continuous population. 3. The values X_i and Y_i represent are ordered (at least the ordinal scale), so the comparisons "greater than", "less than", and "equal to" are meaningful.	Linear Combinations of Means
Multiple Comparisons	**Test Normality:** Shapiro-wilk test, Anderson-Darling test, Kolmogorov–Smirnov Test	

Hypotheses for nonparametric tests

SELECTION OF UNITS	By Randomization	Not by Randomization
At Random	H_0: Each unit in the population would have the same outcome regardless of the treatment assigned; H_a: Treatment effect is non-zero, $\delta \neq 0$, in the population.	H_0: Outcome distributions in the two groups within the population are the same. H_a: There is an association between group status and outcomes in the population. *If similar shapes and spreads :* H_a: Median(X) ≠ Median(Y) (or "≥", or "≤").
Not at Random	H_0: Each unit in the sample would have the same outcome regardless of the treatment assigned; H_a: Treatment effect is non-zero, $\delta \neq 0$, in the sample.	H_0:Outcomes in the sample are not related to group status; H_a: There is an association between group status and outcomes in the sample.

Derive sampling distribution: Q2d on Hw 3
Review "Overview" in lecture 13
Central Limit Theorem
Law of large numbers
Internal vs. external validity

Study design concepts (units, target population, SRS vs. stratified

random sampling, etc.);

} Sample statistics vs. population parameters;

} Rules for expectations and variances of r.v.;

} Mechanics of inference using randomization (i.e., random

permutation) vs. random sampling;

} Tests for 1, 2, and 3+ samples:

} Hypotheses that they are testing (depending on a question of interest,

study design, and data transformation used, if any);

} Assumptions for each test and methods for checking them;

} Test-statistics for each test and their sampling distributions, looking up

p-values using a table, calculating confidence intervals;

} Ability to interpret output of R functions that perform these tests;

} Ability to state conclusions and comment on a scope of inference.

} Definition of distributions by representation;
Conceptual, derivations, other non-simulation questions,

similar to those on homeworks.

} Motivation for and interpretation of simulations on

homeworks.

} Tests and confidence-interval calculation using

distribution tables, comments on assumptions.

} Explanation of what happens behind any computer

function you've used.

} Suggestion: verify the calculations behind t-tests and

other tests for some of the homework problems.

} Ramsey and Schafer has lots of problems to work

through (including conceptual ones)

Study (experimental) unit (subject) - one member of a set of entities being studied.
Parameter (also, estimand) -proportion of childless households in the population.
Estimate - proportion of childless households in the sample.
Parameter is a population characteristic.
Estimand is a parameter that is being estimated (**μ**).
Statistic is a function of the data (therefore, it is random).
Statistic: A function of the data, **μ**= **μ** (y). Test Statistic: Statistic used to weigh evidence supporting and contradicting the null hypothesis.
Reference Distribution: Probability distribution of the test statistic, assuming that the null hypothesis istrue, f(**μ** (Y)|H0).
Estimator is a statistic used as a guess for the value of the estimand ( ).
Estimate is a quantity, which is a particular realization of the estimator (4/12 = 0.33).
Target Population: A collection of units a researcher is interested in; a group about which the researcher wishes to draw conclusions.
Sampling frame: Collection of units that are potential members of the sample.
Overcoverage: selecting units that are outside your sampling frame.
Undercoverage: excluding units that are within your sampling frame.
Sample: A [randomly selected] subset of a sampling frame
Simple Random Sampling (SRS) – every subset of n units has equal chance to be selected
Stratified Sampling – split the population into homogeneous subpopulations and use SRS (or another method) within a sampling frame of each subpopulation.
Systematic Random Sampling - select every k'th unit from the ordered sampling frame, starting randomly from of the first k positions.
Variable probability sampling – allow units to have nonequal probabilities of being sampled.
p-value - probability that the test statistic would be at least as extreme as observed, under the null hypothesis. (or "strength of evidence against the H0")
Significance level (α) is the criterion compared against the p-value. The null hypothesis is rejected if p-value is lower than α.
α reflects the probability of rejecting the null hypothesis given that it is true (**Type I error**).
Scope of inference
Internal validity: are assumptions of the test satisfied?
External validity: possible to make inference to a broader population?
Sampling distribution of a statistic is a (reference) distribution that arises from a chance mechanism used to select a random sample from a population.
Type I Error (or significance level α): Probability of rejecting the null hypothesis, when the null hypothesis is true.
Type II Error (β): Probability of failing to reject the null hypothesis when the null hypothesis is false.
Power (1-β): Probability of rejecting the null hypothesis, when a particular alternative hypothesis is true.
If alpha = 0, beta = 1; if alpha = 1, beta = 0; if alpha increases, beta decreases (power increases).
Increase power by: increase alpha; do one-tailed test; increase effect size; increase sample size
Most to least power, if assumption of t-tools is met
For small samples:
1. T-tools
2. Permutation test
3. Rank-sum test (or signed-rank test for paired data)
4. (Sign test for paired data)
Large samples rejection rates are basically equivalent

Individual confidence level is a probability that a single confidence interval covers the true value.
Familywise confidence level is a probability that all confidence intervals cover the corresponding true values.

Common distributions and calculations:

Population Variance: $\sigma^2 = E\left((Y-\mu)^2\right)$

Sample Variance:

Sample variance is chi-sq distributed: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$

Let Z0~N(0,1), independent of variance X^2~X²ₙ, then $\frac{Z_0}{\sqrt{X^2/n}} \sim t_n$

$$X^2 = \sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$$

Chi-sq distribution with n df:

Pooled sample variance: $S_p^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}$, where

$$S_X^2 = \frac{\sum_{i=1}^{n_X}(X_i - \bar{X})^2}{n_X - 1}, \quad S_Y^2 = \frac{\sum_{i=1}^{n_Y}(Y_i - \bar{Y})^2}{n_Y - 1}$$

Sampling distribution of pooled sample variance: $S_p^2 \sim \frac{\sigma^2}{n_X + n_Y - 2}\chi^2_{n_X + n_Y - 2}$

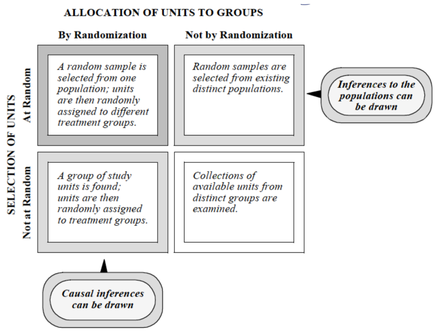
F-distribution: if $X \sim \chi^2_{n_X}$ and $Y \sim \chi^2_{n_Y}$, independent of each other, then

$$\frac{X/n_X}{Y/n_Y} = R \sim F_{n_X, n_Y}$$

(F-distribution with nx and ny d.f)

	H_0 is true	H_0 is false
Reject H_0	α (type I error, significance level)	$1-\beta$ (power)
Do not reject H_0	$1-\alpha$	β (type II error)

Sum of Squared Residuals	General Name	Calculation	Degrees of freedom	Distribu-tion	Corresponding variance (also called "Mean Square")
SSR_{full} Residual sum of squares (full)	SSW (Within Groups)	$\sum_{i=1}^{I}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i)^2$	n-I	$\sigma^2\chi^2_{n-I}$	S_p^2 MSR_{full}
ESS (Extra sum of squares)	SSB (Between Groups)	$\sum_{i=1}^{I} n_i(\bar{Y}_i - \bar{Y})^2 =$ $\sum_{i=1}^{I} n_i(\bar{Y}_i)^2 - n(\bar{Y})^2$	I-1	$\sigma^2\chi^2_{I-1}$	MSB
$SSR_{reduced}$ Residual sum of squares (reduced)	SST (Total)	$\sum_{i=1}^{I}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y})^2$	n-1	$\sigma^2\chi^2_{n-1}$	S^2 Variance of the entire sample as one group



Hypotheses for nonparametric tests

SELECTION OF UNITS	By Randomization	Not by Randomization
At Random	H_0: Each unit in the population would have the same outcome regardless of the treatment assigned; H_a: Treatment effect is non-zero, $\delta \neq 0$, in the population.	H_0: Outcome distributions in the two groups within the population are the same. H_a: There is an association between group status and outcomes in the population. *If similar shapes and spreads :* H_a: Median(*X*) ≠ Median(*Y*) (or "≥", or "≤").
Not at Random	H_0: Each unit in the sample would have the same outcome regardless of the treatment assigned; H_a: Treatment effect is non-zero, $\delta \neq 0$, in the sample.	H_0:Outcomes in the sample are not related to group status; H_a: There is an association between group status and outcomes in the sample.

(Fisher's) Randomization Test: is a distribution-free test for treatment effect in **randomized experiments.**
H0 : Zero treatment effect for all units, δ=0. Each unit's outcome is the same, regardless of the treatment assigned. Consequently, the distribution of outcomes is identical in two groups
Ha : Non-zero treatment effect for ALL units, δ≠0 .
Assumptions:
1. Random assignment to groups.
2. Under the H0, independence of study units.

Test statistic: Difference between average outcomes in the two groups*. $\bar{Y}_c - \bar{Y}_v$
*In fact, any summary of the data can represent a test statistic.
Randomization distribution is a reference distribution of a test statistics in a randomization test, where variation is due to random assignment of the treatment.

Permutation test is a distribution-free nonparametric test for association between group status and outcome in **observational studies.**
H0 : Outcomes not related to group status.
Ha : Outcomes are related to (associated with) group status.
Assumptions: Independence of study units.
Test statistic: Difference in average observed outcomes
between the two groups (or, any other statistic),
Permutation distribution is a reference distribution of a test statistics in a permutation test.
NOTE: Permutation tests are actually evaluating whether the distributions are equivalent, not just the equality of the means:

One Sample z-Test
H0: **μ = μ0**
HA: **μ != μ0**

$$Z = \frac{\bar{Y} - \mu_0}{(\sigma/\sqrt{n})}$$

Test Statistic:
(make sure to double p-value if doing 2-sided test, of ½ alpha)
Exact sampling distribution of Z under H0: Z~N(0,1)
Assumptions:
1. Independence
2. Known population variance, σ^2
3. Normality (CLT allows for deviations!)
4. Random sampling from population

One Sample t-Test
H0: **μ = μ0**
HA: **μ != μ0**

$$T = \frac{\bar{Y} - \mu_0}{(S/\sqrt{n})}$$

Test Statistic:
Exact sampling distribution of T under H0: T ~t_{n-1} (t-distribution with df = n-1)
Assumptions:
1. Independence
2. Normality (CLT allows for deviations!)
3. Random sampling from population

Two Sample z-Test
H0: **μx – μy = μ0**
HA: μx – μy != μ0
Test Statistic: $Z = \frac{(\bar{X} - \bar{Y} - \mu_0)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$ $Z \sim N(0,1)$

Pooled Two-Sample t-Test
H0: μx – μy = μ0
HA: μx – μy != μ0
Test Statistic: $T = \frac{(\bar{X} - \bar{Y} - \mu_0)}{S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}$
Exact sampling distribution of T under H0: T ~$t_{nx+ny-2}$ (t-distribution with df = nx + ny -2)
Note:s $(\bar{X} - \bar{Y}) \sim N\left(\mu_X - \mu_Y, \sigma^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)\right)$,
$\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}$ is an unbiased estimator of $\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}$
Assumptions: same as Welch's, but
If using pooled variance – assume equal variances between populations.

Unpooled (Welch) Two-Sample t-Test
H0: μx – μy = μ0
HA: μx – μy != μ0
Test Statistic: $T = \frac{(\bar{X} - \bar{Y} - \mu_0)}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}$
Approximate Sampling distribution:
$T \dot\sim t_{\tilde{n}}$, where $\tilde{n} = \left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2 / \left(\frac{s_X^4}{n_X^2(n_X - 1)} + \frac{s_Y^4}{n_Y^2(n_Y - 1)}\right)$
Assumptions:
1. Independence between units: (within each population (Yi with Yj and Xi with Xj)) And)between populations (Yi with Xj)
2. Homogeneity of units within each population:(Equal means within each population; Equal variances within each population;
If using pooled variance – assume equal variances between populations.
3. Population distributions are Normal:CLT allows for deviations.
4. Random sampling from populations.

Paired t-Test
H0: μD = μ0
HA: μD != μ0
Test Statistic:

$$X_i - Y_i = D_i \overset{i.i.d.}{\sim} N(\mu_D, \sigma^2) \qquad T = \frac{\bar{D} - \mu_0}{(S_D/\sqrt{n})}$$

Exact sampling distribution of T under H0: T~t_{n-1} (where n is # of pairs)
Assumptions same as one-sample t-test

F-test for equal variance (could also use levene's test): (caution: will reject for large samples when ratio is close to 1)

H_0: $\sigma_X^2 = \sigma_Y^2$

H_A: $\sigma_X^2 \neq \sigma_Y^2 \Leftrightarrow \sigma_X^2 / \sigma_Y^2 \neq 1$

Test Statistic: $F = \frac{S_X^2}{S_Y^2} \sim F_{n_X - 1, n_Y - 1}$

Same assumptions as t-test:

In testing the equality of population variances, two assumptions are required: independent samples and normally distributed populations.

2-sided p-value is just 2*the one-sided value.

Test Normality:
Shapiro-wilk test, Anderson-Darling test, Kolmogorov–Smirnov Test

How does CLT apply?
 # checking assumptions:
Independence: plot subgroups, and see if relationship still holds (cluster effect)
Plot data over time (serial effects)
Plot data vs space (spatial effect)
Normality: Histograms should look normal – overlay with normal curve and kernel density plot
Boxplot or Q-Q plot will show outliers
Equal Variance: divide variances to see if ratio is <0.5 or > 2

T-test is robust:
-if populations are symmetrically skewed
-if sample sizes are equal (and large)
-when sample sizes are equal, pooled-ttest is robust to unequal variances
Sensitive: When the sample sizes are not equal, t-tests are more sensitive to skewedness and long-tailedness.
For small samples, t-tests are somewhat sensitive to markedly different skewedness in two groups.
Watch out for outliers.

Robustness of F-tests:
Not resistant to outliers. Fairly robust to non-normality. Independence w/I and b/w groups is necessary. Equal variance is necessary.

F-Test Assumptions:
Equal variances is crucial
Not resistant to outliers
Normality is not critical.

How does CLT apply?
 # checking assumptions:
Independence: plot subgroups, and see if relationship still holds (cluster effect)
Plot data over time (serial effects)
Plot data vs space (spatial effect)
Normality: Histograms should look normal – overlay with normal curve and kernel density plot
Boxplot or Q-Q plot will show outliers
Equal Variance: divide variances to see if ratio is <0.5 or > 2

T-test is robust:
-if populations are symmetrically skewed
-if sample sizes are equal (and large)
-when sample sizes are equal, pooled-ttest is robust to unequal variances
Sensitive: When the sample sizes are not equal, t-tests are more sensitive to skewedness and long-tailedness.
For small samples, t-tests are somewhat sensitive to markedly different skewedness in two groups.
Watch out for outliers.

Robustness of F-tests:
Not resistant to outliers. Fairly robust to non-normality. Independence w/I and b/w groups is necessary. Equal variance is necessary.

F-Test Assumptions:
Equal variances is crucial
Not resistant to outliers
Normality is not critical.

Confidence intervals: Yi ~iid N(mu, sigma^2)

One Sample t-test:

$$\left(\bar{Y} - t_{1-\alpha/2, n-1} \frac{S}{\sqrt{n}}, \quad \bar{Y} + t_{1-\alpha/2, n-1} \frac{S}{\sqrt{n}} \right)$$

Two Sample t-test:

- Two-sample *t*-tests

Standard Error

$$\left(\bar{X} - \bar{Y} - t_{1-\alpha/2, df} \cdot SE, \quad \bar{X} - \bar{Y} + t_{1-\alpha/2, df} \cdot SE \right)$$

- Pooled *t*-test: $SE = S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}, \; df = n_X + n_Y - 2$

- Unpooled *t*-test: $SE = \sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}, \; df = \tilde{n}$

$$\tilde{n} = \left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y} \right)^2 \Big/ \left(\frac{s_X^4}{n_X^2 (n_X - 1)} + \frac{s_Y^4}{n_Y^2 (n_Y - 1)} \right)$$

Two-Sample t-test for I samples (I > 2) Conf int

- Random sampling.

Pooled using *I* samples!

Confidence intervals: $SE = S_p \sqrt{\frac{1}{n_m} + \frac{1}{n_k}}, \; df = n_1 + n_2 + \ldots + n_I - I$

$$\left(\bar{Y}_m - \bar{Y}_k - t_{1-\alpha/2, df} \cdot SE, \quad \bar{Y}_m - \bar{Y}_k + t_{1-\alpha/2, df} \cdot SE \right)$$

Transformations:
Log Transformation:

1. **Randomized Experiments** (T=Treatment, C = Control)

$$Z_{T,i} = \ln(Y_{T,i}), \; i = 1, \ldots, n_T$$

$$Z_{C,i} = \ln(Y_{C,i}), \; i = 1, \ldots, n_C$$

Responses:

$$Z_{T,i} = Z_{C,i} + \delta \Leftrightarrow \frac{Y_{T,i}}{Y_{C,i}} = e^{\delta}$$

$$\exp\left(\bar{Z}_T - \bar{Z}_C\right) \text{ estimates } \frac{Y_{T,i}}{Y_{C,i}}$$

Interpretation: -- The response of an experimental unit to the treatment will be exp(ZbarT – ZbarC) times as large as its response to the control

2. Observational studies with **random sampling**:

$$\exp\left(\bar{Z}_2 - \bar{Z}_1\right) \text{ estimates } \frac{m_2}{m_1}$$

Interpretation: -- The median of the second population is exp(Zbar2 – Zbar1) times as large as the median of the first population.

3. In paired t-test: $Z_i = \ln(Y_{2,i}) - \ln(Y_{1,i}), \; i = 1, \ldots, n$

Randomized Experiments:

$\exp(\bar{Z})$ estimates a multiplicative treatment effect $\frac{Y_{2,i}}{Y_{1,i}}$

Observation studies w/ rand. Sampling:

$$\text{Median}(Y_{2,i} / Y_{1,i}) = m, \text{ then } \exp(\bar{Z}) \text{ estimates } m.$$

Square Root Transformation sqrt(Y): Good for moderately skewed data, but hard to interpret results
Reciprocal Transformation 1/Y: good for severely skewed data, can be used with negative data
Logit log(Y/(1-Y)) or arcsin(2Y-1): good for proportions

Nonparametric alternatives to t-tools:

Rank Sum test:
Hypotheses:
H0: (if treatments are randomized): Same as fisher's randomization test
H0: (if treatments are not randomized): same as permutation test
* If shapes and spreads of the two populations are similar: HA: There is a difference in medians between the two populations.
Transform data:

$$Z_{1,i} = \text{Rank}(X_i \mid X_1, \ldots, X_{n_X}, Y_1, \ldots, Y_{n_Y}), \; i = 1, \ldots, n_X$$

$$Z_{2,i} = \text{Rank}(Y_i \mid X_1, \ldots, X_{n_X}, Y_1, \ldots, Y_{n_Y}), \; i = 1, \ldots, n_Y$$

Ties are averaged

Test statistic: $T = \sum_{i=1}^{n_X} Z_{1,i}$, assuming $n_X \leq n_Y$

Assumptions: Independence

Normal approximation to rank Sum test (if no ties, and samples greater than 10):

$$T \sim N\left(\frac{n_X (n_X + n_Y + 1)}{2}, \frac{n_X n_Y (n_X + n_Y + 1)}{12} \right)$$

With ties, Rbar is sample mean (mean of all ranks) and Sr^2 is sample variance for combined set of nx + ny ranks:

$$T \,\dot{\sim}\, N\left(n_X \bar{R}, S_R^2 \frac{n_X n_Y}{n_X + n_Y} \right)$$

We can use this Z-statistic:

$$Z = \frac{T - n_X \bar{R}}{S_R \sqrt{\frac{n_X n_Y}{n_X + n_Y}}}$$

Sign test for paired data (lower power than Wilcoxon signed-rank test)

General formuation: H_0: $P(D_i > 0) = P(Y_i > X_i) = 0.5$, H_a: $P(D_i > 0) \neq 0.5$ (or >0.5, or <0.5).

If differences are continuous rv's: H_0: Median$(D_i) = 0$, H_a: Median$(D_i) \neq 0$ (or >0, or <0).

If differences are symmetric: H_0: $E(D_i) = 0$, H_a: $E(D_i) \neq 0$ (or >0, or <0)

Test Statistic: K = number of positive differences.
Exact distribution of K: K ~Binom(m, 0.5), where m is final number of pairs with nonzero Di.

Normal approximation (for large samples): $Z = \frac{K - m/2}{\sqrt{m/4}} \,\dot{\sim}\, N(0,1)$

Assumptions:

4. The differences D_i are assumed to be independent.
5. Each D_i comes from the same continuous population.
6. The values X_i and Y_i represent are ordered (at least the ordinal scale), so the comparisons "greater than", "less than", and "equal to" are meaningful.

Wilcoxon Signed-Rank Test for Paired Data:
(more assumptions, but higher power than sign test)
Hypotheses: General formulation:
H0: The rank of the magnitude of within-pair difference is unrelated to the sign of the difference.
Ha: The rank of the magnitude of within-pair difference is related to the sign of the difference.

If differences are continuous r.v.'s:
H0: Median(Di)=0
Ha: Median(Di)≠0 (or >0, or <0)

If the distribution of differences is symmetric :
H0: E(Di)=0
Ha: E(Di)≠0 (or >0, or <0)

1. Calculate differences (+ or -).
2. Rank the absolute values of differences |Di|
3. $Z_i = \text{Rank}(|D_i|),\ i = 1,\dots,n.$

Test Statistic: (sum of ranks for positive differences)

$$S = \sum_{i=1}^{n} Z_{i|D_i>0}$$

Exact sampling distribution: do permutation test, and switch group status for each pair.
Normal approximation (for large sample sizes, m is >= 20, where m is the number of non-zero differences):

$$Z = \frac{S - m(m+1)/4}{\sqrt{m(m+1)(2m+1)/24}} \sim N(0,1)$$

Assumptions:
1. Data are paired and come from the same population.
2. Each pair is chosen randomly and independently.
3. The data are measured at least on an ordinal scale, but need not be normal.

Kruskal-Wallis Test (Nonparametric ANOVA)
H0: Median(Y1) = Median(Y2) = …=Median(YI)
HA: at least one median is different from the others

Test Statistic – convert to ranks, and do ANOVA on ranks:

$$KW = \frac{\text{ESS (on ranks)}}{S_R^2} \overset{H_0}{\sim} \chi^2_{I-1}$$

S_R^2 is the sample variance of **all** ranks (not pooled)
Assumptions:
1. Independence.
2. Random Sampling

Two-Sample t-Test for I samples
H0: μm – μk = μ0
HA: μm – μk != μ0

Test Statistic:

$$T = \frac{(\bar{Y}_m - \bar{Y}_k - \mu_0)}{S_p\sqrt{\frac{1}{n_m} + \frac{1}{n_k}}};\quad S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + \dots + (n_I-1)S_I^2}{n_1+n_2+\dots+n_I - I}$$

Exact sampling distribution of T under H0 is:
$T \sim t_{n1+n2+\dots+nI-I}$, where I is number of groups.
Assumptions:
Independence of units within each group;
Independence of units between groups;
Normality of all I populations;
Equal variances for all I populations;
Random sampling.

Anova F-test for Equality of All means in I samples
H0: μ1 = μ2 = …=μi (reduced model – equal means model)
Ha: at least one mean is different from the others (full model – separate means model)

SSRfull = SSW (within groups)

$$S_p^2 = \frac{\text{SSR}_{\text{full}}}{n-I}$$

$$S_p^2 \sim \frac{\sigma^2}{n-I}\chi^2_{n-I} \Rightarrow \text{SSR}_{\text{full}} \sim \sigma^2\chi^2_{n-I}, \text{ d.f.} = n\text{-}I$$

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + \dots + (n_I-1)S_I^2}{n-I}$$

SSRreduced = SST (**total error**)
S^2 is the sample variance of the entire sample, taken as one group:

$$S^2 = \frac{\text{SSR}_{\text{reduced}}}{n-1}$$

$$S^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1} \Rightarrow \text{SSR}_{\text{reduced}} \sim \sigma^2\chi^2_{n-1}, \text{ d.f.} = n\text{-}1$$

ESS = SSB (between groups)
ESS = SSRreduced - SSRfull

$$\text{ESS} \overset{H_0}{\sim} \sigma^2\chi^2_{I-1}, \text{ d.f.} = I\text{-}1$$

Test Statistic:

$$R = \frac{\text{ESS}/(I-1)}{\text{SSR}_{\text{full}}/(n-I)} = \frac{\text{MSB}}{\text{MSR}_{\text{full}}}$$

MSB = Mean square between
MSRfull=Mean square residuals (within groups)

Easy way to calculate: MSRfull == S_p^2
Exact sampling distribution of R under H0:

$$F_{(I-1,n-I)}$$

Assumptions (check w/ Q-Q or residual plot):
1. Normality of populations;
2. Equal population variances for all groups;
3. Independence within each group;
4. Independence between groups;
5. Homogeneity within each group;
6. Random sampling

F-test for Equality in a subset of groups:
"Are groups 2-I all the same?"

$H_0: \mu_1$ and $\mu_2 = \mu_3 = \dots = \mu_I$ (Reduced Model)

$H_A: \mu_1,\ \mu_2,\ \mu_3,\ \dots,\ \mu_I$ (Full Model)

Reduced Model: H_0: Equality of means in k_0 groups of I samples
Full Model: H_A: Equality of means in k_A groups of I samples, $k_A > k_0$, such that groups for H_A are nested within H_0 groups.

Test statistic:

$$R = \frac{\text{ESS}/(k_A - k_0)}{\text{SSR}_{\text{full}}/(n - k_A)} \sim F_{(k_A - k_0,\, n-k_A)}$$

Sum of Squared Residuals	General Name	Degrees of freedom	Distribution	Corresponding variance (also called "Mean Square")
SSR_{full} Residual sum of squares (full)	SSW (Within Groups)	$n-k_A$	$\sigma^2\chi^2_{n-k_A}$	MSR_{full} Pooled variance for k_A groups.
ESS Extra sum of squares	SSB (Between Groups)	k_A-k_0	$\sigma^2\chi^2_{k_A-k_0}$	MSB
$SSR_{reduced}$ Residual sum of squares (reduced)	SST (Total)	$n-k_0$	$\sigma^2\chi^2_{n-k_0}$	$MSR_{reduced}$ Pooled variance for k_0 groups.

ANOTHER VERSION: F-test for Equality in a subset of groups:
"Is group 1 different from all other groups?"

$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_I = \mu$ (Reduced Model)

$H_A: \mu_1,$ vs. $\mu_2 = \mu_3 = \mu_I = \mu_{Other}$ (Full Model)

Note: SSRfull now is equal to SSRreduced from above

Hypothesis test for linear combinations of means:
"Is one group mean equal to the average of a combination of other group means?"

A linear combination of group means is : $\gamma = C_1\mu_1 + C_2\mu_2 + \dots + C_I\mu_I$
, where Ci's are fixed coefficients chosen by the researcher

$$H_0: \gamma = \gamma_0$$
$$H_A: \gamma \neq \gamma_0$$

If all coefficients sum to 0, then the linear combination is called **a contrast**.
Estimate of linear combination:

$$G = C_1\bar{Y}_1 + C_2\bar{Y}_2 + \dots + C_I\bar{Y}_I$$

Sampling distribution of G under H0:

$$G \sim N\left(\gamma_0, \sigma^2\left(\frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \dots + \frac{C_I^2}{n_I}\right)\right)$$

T-test for Linear combination of means:

$$T = \frac{g - \gamma_0}{SE(g)} = \frac{C_1\bar{Y}_1 + C_2\bar{Y}_2 + \dots + C_I\bar{Y}_I - \gamma_0}{S_p\sqrt{\frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \dots + \frac{C_I^2}{n_I}}}$$

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + \dots + (n_I-1)S_I^2}{n-I}$$

(same as F-test s-pooled)

Approximate sampling distribution: $T \sim t_{n_1+n_2+\dots+n_I-I}$
Assumption: Same as ANOVA F-test

Mulitple Comparisons:
Familywise Type-I error

$$\alpha \le 1-(1-\alpha)^N \le N\alpha$$

N perfectly dependent tests <= probability of at least one type-I error among N indep. Tests <= max familywise error

Bonferroni: (When all else fails)
Set individual significance levels at alpha/N, where N is the # of tests
For pairwise mean comparisons, half width of CI
= $t_{n-I,\ 1-\alpha/(2N)}SE$

Tukey HSD (Honest Significant Difference): (when interested in all pairs of groups)
- Considers the largest difference between any two sample means for I groups
- ensures familywise conf leve = 1-alpha

$$Q = \frac{\bar{Y}_{\max} - \bar{Y}_{\min}}{S_p/\sqrt{\bar{n}}} \sim \boxed{q(I, n-I)},\ \text{where } n = \bar{n}I$$

Studentized Range Distribution

, nbar is # in each sample), I is # of groups
Half –width for CI:

$$\frac{q(I, n-I, 1-\alpha)}{\sqrt{2}}SE, \text{ where } SE = S_p\sqrt{2/\bar{n}}$$

Assumptions:
Normality, equal variances, equal sample sizes

Tukey-Kramer Procedure (like Tukey HSD, but no need for equal sample sizes):

$$\frac{q(I, n-I, 1-\alpha)}{\sqrt{2}}SE, \text{ where } SE = S_p\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

Formulas:

N Choose k: $\frac{n!}{k!(n-k)!}$

Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p-value
Between Groups	1,927.08	6	321.18	6.72	.000061
Within Groups	1,864.45	39	47.81		
Total	3,791.53	45			

1 Sum of squared residuals from fitting the full (separate-means) model
2 Sum of squared residuals from fitting the reduced (equal-means) model
3 degrees of freedom
n-I
n-1
4 Subtract the "Within" from the "Total "
5 A mean square is the ratio of a sum-of-squares to its degrees of freedom
6 The F-statistic is the ratio of the Between MS to the Within MS
7 The p-value comes from an F-distribution with 6 and 39 df
NOTE: *This is* s_p^2

Log-transformed, observational t-test

```
Two Sample t-test

data:  log(Salary) by Sex
t = -6.1715, df = 91, p-value = 1.849e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.19423950 -0.09964777
sample estimates:
mean in group Female   mean in group Male
            8.539048             8.685992
```

What is the 95% confidence interval for the ratio of population median salaries?

Solution: *The 95% CI for the ratio is* $(e^{-0.194}, e^{-0.0996}) = (0.824, 0.905)$.

When does pooled t-test rejection rate differ:
If the variances are actually different, the pooled t-test may be invalid. If the larger variance comes from the group with smaller sample size, the t- test tends to reject the null too often. (One point was taken off if you didn't explain that the rejection rate tends to be too high when the larger variance is in the group with the smaller sample size; the rejection rate tends to be too low when the larger variance is in the group with the larger sample size).

Derive the mean and variance of the signed-rank test statistic under its null hypothesis in a study with N pairs. Show your work/explain your answer: just writing down the correct mean and variance is not sufficient for full credit. [6 points]

Answer:

If there are N pairs, for each pair i, denote $D_i = 1$ if 1st observation is larger, and $D_i = 0$ if 2nd observation is larger.

$$D_i \sim Ber(1/2)\ i.i.d.$$

Let T=test stat=$\sum_{i=1}^{N} D_i * i$.

$$E(T) = E(\sum_{i=1}^{N} D_i * i) = \sum_{i=1}^{N} i * E(D_i) = \frac{1}{2}\sum_{i=1}^{N} i = \frac{N(N+1)}{4}.$$

Because all the D_is are independent of each other,

$$Var(T) = Var(\sum_{i=1}^{N} D_i * i) = \sum_{i=1}^{N} i^2 * Var(D_i) = \frac{1}{4}\sum_{i=1}^{N} i^2 = \frac{N(N+1)(2N+1)}{24}.$$

$$\frac{\bar{X} - \mu_X}{\frac{\sigma}{\sqrt{n_X}}} \sim N(0,1),\ \frac{S_p^2}{\sigma^2} \sim \frac{1}{n_X + n_Y - 2}\chi^2_{n_X+n_Y-2}.$$

$\frac{\bar{X}-\mu_X}{\frac{\sigma}{\sqrt{n_X}}}$ and $\frac{S_p^2}{\sigma^2}$ are independent. By definition, $t_n \sim \frac{N(0,1)}{\sqrt{\chi^2/n}}$, where the numerator and the denominator are independent. Therefore,

$$\frac{\bar{X} - \mu_X}{S_p/\sqrt{n_X}} = \frac{\bar{X} - \mu_X}{\frac{\sigma}{\sqrt{n_X}}} \frac{1}{\sqrt{\frac{S_p^2}{\sigma^2}}} \sim t_{n_X+n_Y-2}.$$

Derive test stat for paired t-test

Name	Param.	PMF or PDF	Mean	Variance
Bernoulli	p	$P(X=1)=p, P(X=0)=q$	p	pq
Binomial	n, p	$\binom{n}{k}p^k q^{n-k}$, for $k \in \{0,1,\dots,n\}$	np	npq
FS	p	pq^{k-1}, for $k \in \{1,2,\dots\}$	$1/p$	q/p^2
Geom	p	pq^k, for $k \in \{0,1,2,\dots\}$	q/p	q/p^2
NBinom	r, p	$\binom{r+n-1}{r-1}p^r q^n, n \in \{0,1,2,\dots\}$	rq/p	rq/p^2
HGeom	w, b, n	$\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$, for $k \in \{0,1,\dots,n\}$	$\mu = \frac{nw}{w+b}$	$(\frac{w+b-n}{w+b-1})n\frac{\mu}{n}(1-\frac{\mu}{n})$
Poisson	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$, for $k \in \{0,1,2,\dots\}$	λ	λ
Uniform	$a < b$	$\frac{1}{b-a}$, for $x \in (a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	μ, σ^2	$\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2
Expo	λ	$\lambda e^{-\lambda x}$, for $x > 0$	$1/\lambda$	$1/\lambda^2$
Gamma	a, λ	$\Gamma(a)^{-1}(\lambda x)^a e^{-\lambda x}x^{-1}$, for $x > 0$	a/λ	a/λ^2
Beta	a, b	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$, for $0 < x < 1$	$\mu = \frac{a}{a+b}$	$\frac{\mu(1-\mu)}{a+b+1}$
χ^2	n	$\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}$, for $x > 0$	n	$2n$
Student-t	n	$\frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$	0 if $n > 1$	$\frac{n}{n-2}$ if $n > 2$

The function Γ is given by $\Gamma(a) = \int_0^\infty x^a e^{-x}\frac{dx}{x}$ for all $a > 0$. For any $a > 0$, $\Gamma(a+1) = a\Gamma(a)$. We have $\Gamma(n) = (n-1)!$ for n a positive integer, and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Derive distribution for T^2

(5 pts) Let $T \sim t_{n-1}$. Derive the distribution of T^2.

(2 pts for a precise, neat RV set-up) Let us define two independent random variables Z and S^2, where $X \sim N(0,1)$, and $S^2 \sim \chi^2_{n-1}$. Then, using the definition by representation, $Z^2 \sim \chi^2_1$. The representation of a random variable $T \sim t_{n-1}$ is

$$T = \frac{Z}{S/\sqrt{n-1}}$$

$$(3\ pts) \Rightarrow T^2 = \frac{Z^2}{S^2/(n-1)} \sim \frac{\chi^2_1/1}{\chi^2_{n-1}/(n-1)} \sim F_{1,n-1}$$

because Z^2 and S^2 are independent.