

Biostat 200

Lecture 7

Outline for today

- Hypothesis tests so far
 - One mean, one proportion, 2 means, 2 proportions
- Comparison of means of multiple independent samples (ANOVA)
- Non parametric tests
 - For paired data
 - For 2 independent samples
 - For multiple independent samples

Hypothesis tests so far

Dichotomous data

- Test of one proportion:

Null hypothesis $p=p_0$ (two-sided)

Test statistic $z = (\hat{p} - p_0) / \sqrt{(p_0(1 - p_0)/n)}$

- Proportion test for two independent samples

Null hypothesis $p_1=p_2$ (two-sided)

Test statistic

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1 - \hat{p})[1/n_1 + 1/n_2]}}$$

$$\text{where } \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Hypothesis tests so far

Numerical data

- T-test of one mean:

Null hypothesis: $\mu = \mu_0$ (two-sided)

Test statistic $t = (\bar{X} - \mu_0) / (s / \sqrt{n})$

$n-1$ degrees of freedom

- Paired t-test

Null hypothesis $\mu_1 = \mu_2$ (two-sided)

Test statistic $t = \bar{d} / (s_d / \sqrt{n})$

where $s_d = \sqrt{(\sum (d_i - \bar{d})^2 / (n-1))}$

Stata calculations for PAIRED ttest

When you have the data in the form of the difference between the pairs:

`ttest diffvar==0` e.g. `ttest peth_diff==0`

When you have the data in 2 different variables

`ttest var1==var2` e.g. `ttest peth0_log10==peth3_log10`

When you don't have the data

`ttesti npairs meandiff sddiff nullvalue`

```
ttesti 177 -.101553 .5782935 0
```

One-sample t test

		Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]

	+					

x		177	-.101553	.0434672	.5782935	-.187337 -.015769

mean = mean(x)

t = -2.3363

Ho: mean = 0

degrees of freedom = 176

Ha: mean < 0

Pr(T < t) = 0.0103

Ha: mean != 0

Pr(|T| > |t|) = 0.0206

Ha: mean > 0

Pr(T > t) = 0.9897

Hypothesis tests so far

Numerical data

- Independent samples t-test

Null hypothesis $\mu_1 = \mu_2$ (two-sided)

Test statistic $t = (\bar{x}_1 - \bar{x}_2) / SE(\text{diff between means})$

SE and degrees of freedom depend on assumption of equal or unequal variances

Comparing the t-tests

Assuming equal variance:

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{(s_p^2/n_1 + s_p^2/n_2)}} \text{ where}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Example:

N1=20 s1²=3

N2=40 s2²=5

$$S_p^2 = (19 \cdot 3 + 39 \cdot 5) / 59 = 4.27$$

$$T_{stat \text{ denominator}} = \sqrt{4.27/20 + 4.27/40} = 0.267$$

$$T_{stat \text{ denom, unequal}} = \sqrt{3/20 + 5/40} = 0.524$$

Not assuming equal variance:

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{(s_1^2/n_1 + s_2^2/n_2)}}$$

T-test: equal or unequal variance?

- Why can't we just do a test to see if the variances in the groups are equal, to decide which t-test to use?
 - “It is generally unwise to decide whether to perform one statistical test on the basis of the outcome of another”
 - The reason has to do with Type I error (multiple comparisons, discussed next slide)
 - You are better off always assuming unequal variance if your data are approximately normal

Statistical hypothesis tests

Data and comparison type	Alternative hypotheses	Test and Stata command
Numerical; One mean	$H_a: \mu \neq \mu_a$ (two-sided) $H_a: \mu > \mu_a$ or $\mu < \mu_a$ (one-sided)	Z or t-test • <code>ttest var1=hypoth val.*</code>
Numerical; Two means, <u>paired data</u>	$H_a: \mu_1 \neq \mu_2$ (two-sided) $H_a: \mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	Paired t-test • <code>ttest var1=var2*</code>
Numerical; Two means, independent data	$H_a: \mu_1 \neq \mu_2$ (two-sided) $H_a: \mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	T-test (equal or unequal variance) • <code>ttest var1, by(byvar) unequal</code>
Numerical; Two or more means, independent data		
Dichotomous; One proportion	$H_a: p \neq p_a$ (two-sided) $H_a: p > p_a$ or $p < p_a$ (one-sided)	Proportion test • <code>prtest var1=hypoth value*</code> • <code>bitest var1=hypoth value</code>
Dichotomous; two proportions	$H_a: p_1 \neq p_2$ (two-sided) $H_a: p_1 > p_2$ (one-sided)	Proportion test (z-test) • <code>prtest var1, by(byvar)</code>
Categorical by categorical (nxk)		

Comparison of several means

- The extension of the t-test to several independent groups is called analysis of variance or ANOVA
- Why is it called analysis of variance?
 - Even though your hypothesis is about the means, the test actually compares the variability between groups to the variability within groups

Analysis of variance

The null hypothesis is:

H_0 : all equal means $\mu_1 = \mu_2 = \mu_3 = \dots$

The alternative H_A is that at least one of the means differs from the others

Analysis of variance

- Why can't we just do t-tests on the pairs of means?
 - Multiple comparison problem
 - What is the probability that you will incorrectly reject H_0 at least once *when you run n independent tests* (H_0 is true), when the probability of incorrectly rejecting the null on each test is 0.05?



Analysis of variance

- This is $P(X \geq 1)$ with $p=0.05$, n =number of tests
- X =the number of times the null is incorrectly rejected
- $P(X \geq 1) = 1 - P(X=0)$
- $P(X=0) = \binom{n}{0} (.05)^0 (1-.05)^{n-0}$
- For $n=4$

```
di 1 - (1 - .05)^4
```

```
.18549375
```
- Using the binomial

```
di binomialtail(4, 1, .05)
```

```
.18549375
```

Comparison of several means: analysis of variance

- We calculate the ratio of:
 - The between group variability
 - The variability of the sample means around the overall (or grand) mean $\bar{x}$
 - to the overall within group variability

$$F = \frac{s_B^2}{s_W^2}$$

Between group variability

The between group variability is the variability around the overall (or grand) mean $\bar{x}$

$$s_B^2 = \frac{n_1(\bar{x}_1 - \bar{X})^2 + n_2(\bar{x}_2 - \bar{X})^2 + \dots + n_k(\bar{x}_k - \bar{X})^2}{k - 1}$$

k = the number of groups being compared

n_1, n_2, n_k = the number of observations in each group

$\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ are the group means

$\bar{X}$ = the grand mean – the mean of all the data combined

Within group variability

The within group variability is a weighted average of the sample variances within each group

$$s_W^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{n_1 + n_2 + \dots + n_k - k}$$

k= the number of groups being compared

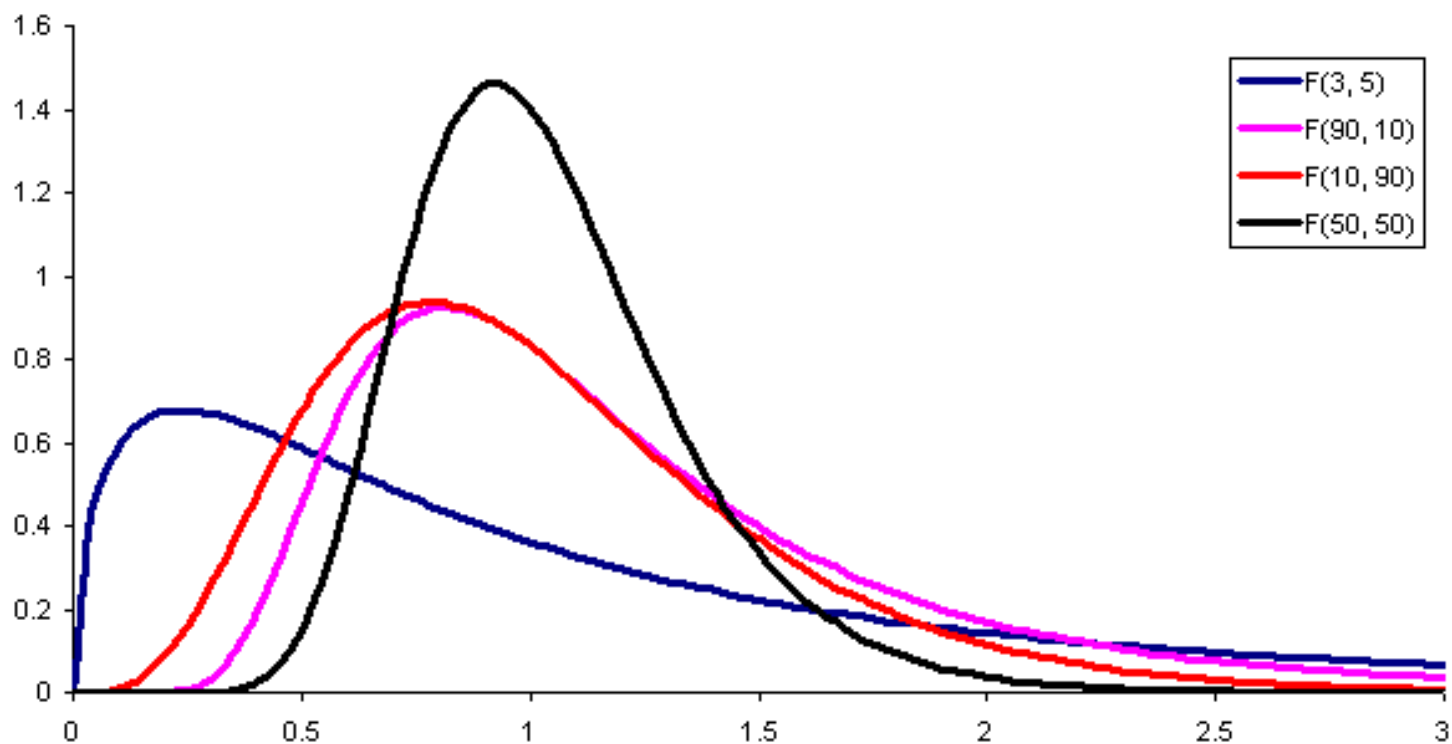
n_1, n_2, n_k = the number of observations in each group

$s_1^2, s_2^2, \dots, s_k^2$ are the sample variances in each group

Comparison of several means: analysis of variance

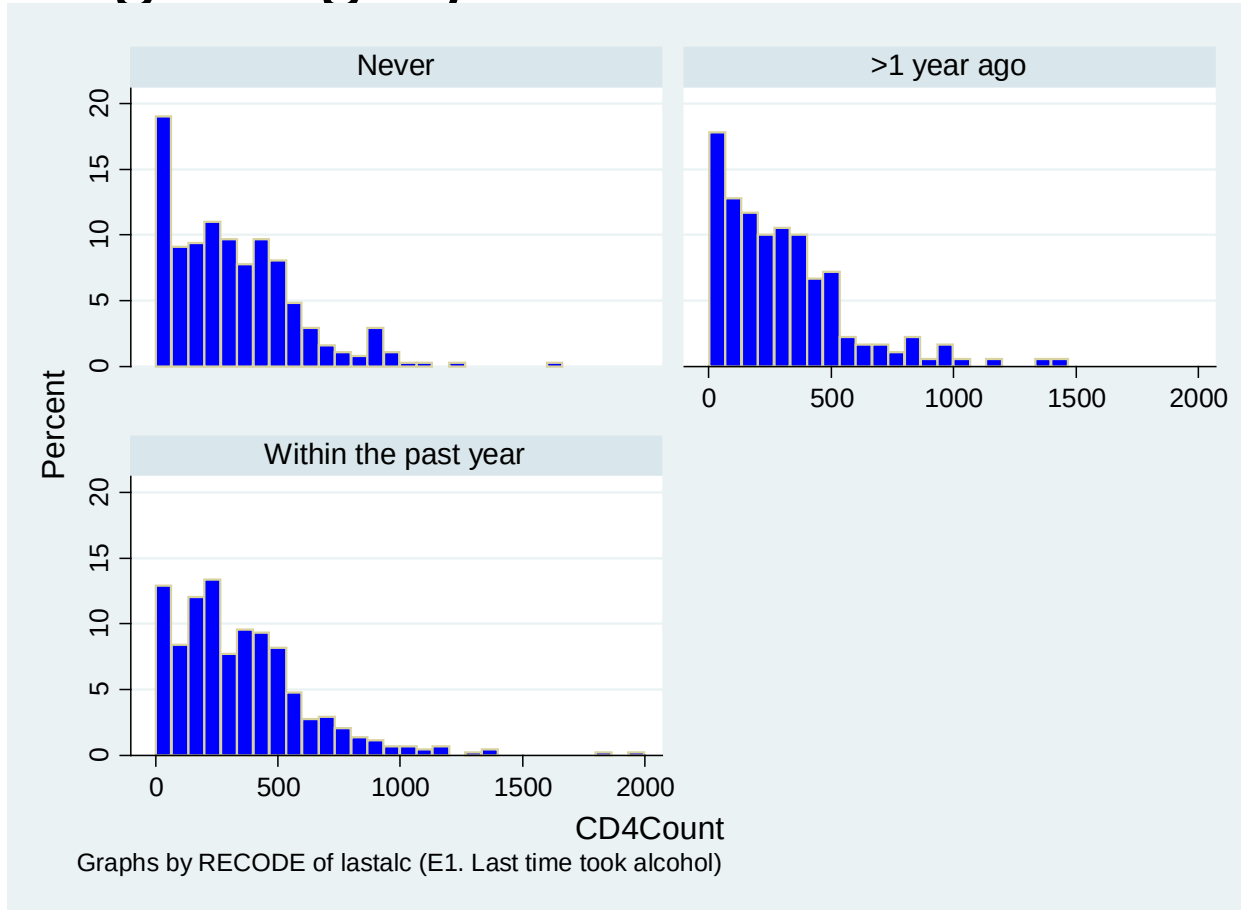
- The test statistic is
$$F = \frac{S_B^2}{S_W^2}$$
- We compare the F statistic to the F-distribution, with k-1 and n-k degrees of freedom
 - k=the number of groups being compared
 - n=the total number of observations

F-distribution

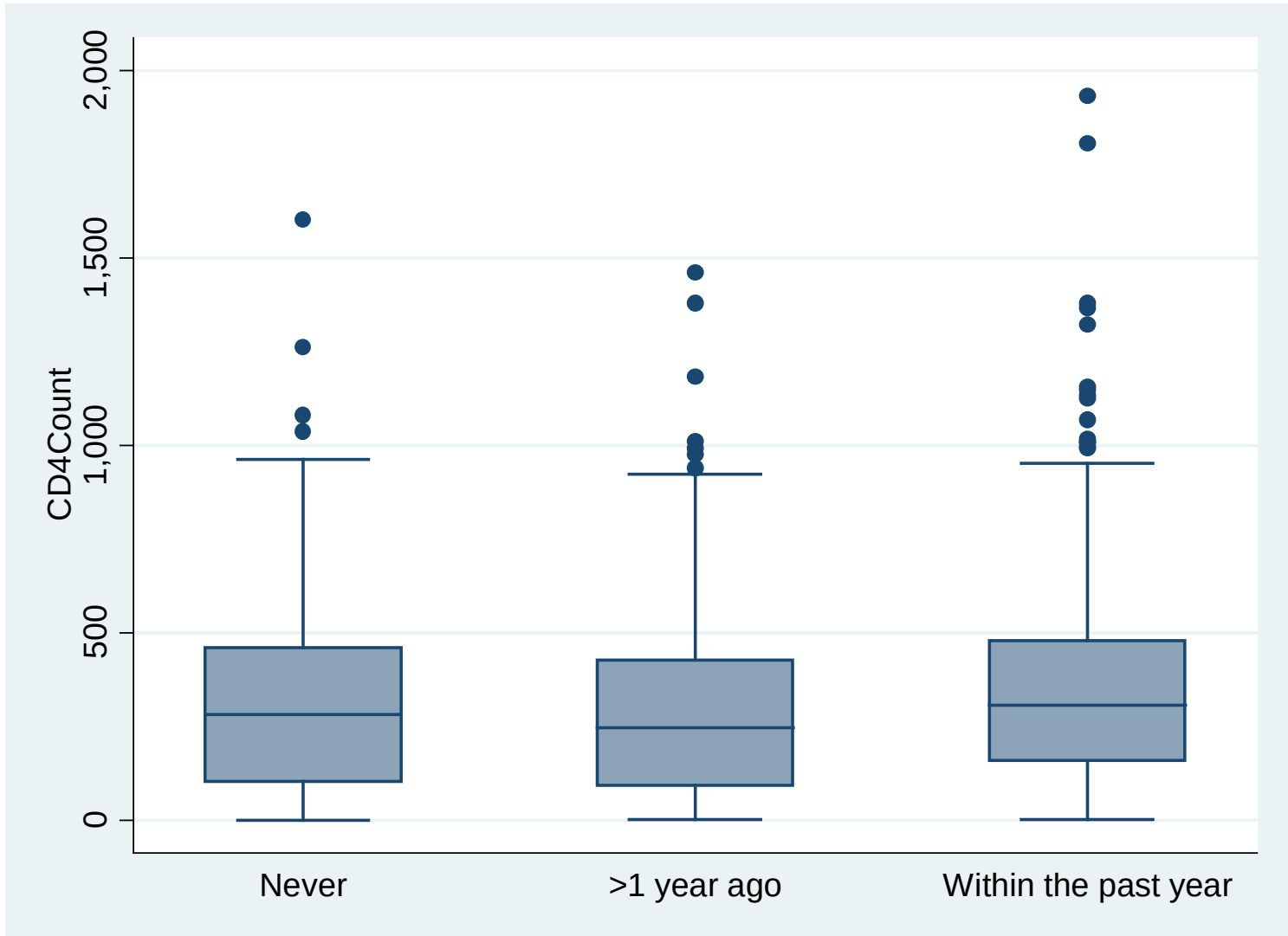


ANOVA example

- Does CD4 count at time of testing differ by drinking category?



```
*Using vct_baseline_biostat200_v1.dta **  
hist cd4count, by(lastalc_3) percent fcolor(blue)
```



```
graph box cd4count, over(lastalc_3)
```

ANOVA example

```
tabstat cd4count, by(lastalc_3) s(n mean sd min median max)
```

Summary for variables: cd4count

by categories of: lastalc_3 (RECODE of lastalc (E1. Last time took alcohol))

lastalc_3	N	mean	sd	min	p50	max
Never	373	317.1475	253.4013	1	283	1601
>1 year ago	180	305.3778	266.9453	2	248.5	1461
Within the past year	441	349.8662	273.9364	3	308	1932
Total	994	329.5322	265.5157	1	285	1932



ANOVA example

- CD4 count, by alcohol consumption category

oneway var groupvar

. oneway cd4count lastalc_3

$$F = \frac{S_B^2}{S_W^2}$$

Analysis of Variance					
Source	SS	df	MS	F	Prob > F
Between groups	344571.162	2	172285.581	2.45	0.0867
Within groups	69660550.3	991	70293.189		
Total	70005121.5	993	70498.6118		

Bartlett's test for equal variances: chi2(2) = 2.4514 Prob>chi2 = 0.294

k=3 groups, n=994 total observations. n-k=991

. di Ftail(2, 991, 2.45)

.08681613

ANOVA example

$$s_B^2 = \frac{n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \dots + n_k(\bar{x}_k - \bar{x})^2}{k-1}$$

```
. oneway cd4count lastalc_3
```

Analysis of Variance					
Source	SS	df	MS	F	Prob > F
Between groups	344571.162	2	172285.581	2.45	0.0867
Within groups	69660550.3	991	70293.189		
Total	70005121.5	993	70498.6118		

Bartlett's test for equal variances: $\chi^2(2) = 2.4514$ Prob> $\chi^2 = 0.294$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

$$s_W^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{n_1 + n_2 + \dots + n_k - k}$$

This is the sample variance if you lumped all the CD4 counts together

ANOVA

- Note that if you only have two groups, you will reach the same conclusion running an ANOVA as you would with a t-test
- The test statistic F_{stat} will equal $(t_{\text{stat}})^2$

T-test vs. F test (ANOVA) example

```
. oneway cd4count sex
```

Analysis of Variance					
Source	SS	df	MS	F	Prob > F
Between groups	521674.035	1	521674.035	7.41	0.0066
Within groups	70155332.6	997	70366.4319		
Total	70677006.7	998	70818.6439		

Bartlett's test for equal variances: $\chi^2(1) = 0.0472$ Prob> $\chi^2 = 0.828$

T-test vs. F test (ANOVA) example

```
. ttest cd4count, by(sex)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
1	374	299.6925	13.6301	263.5935	272.891	326.494
2	625	346.9104	10.65047	266.2618	325.9953	367.8255
combined	999	329.2332	8.419592	266.1177	312.7111	345.7554
diff		-47.21789	17.34162		-81.24815	-13.18762

diff = mean(1) - mean(2)
Ho: diff = 0

degrees of freedom = 997

t = -2.7228

Ha: diff < 0
Pr(T < t) = 0.0033

Ha: diff != 0
Pr(|T| > |t|) = 0.0066

Ha: diff > 0
Pr(T > t) = 0.9967

```
. di 2.7228^2  
7.4136398
```

Multiple comparisons

- If we reject H_0 , we might want to know which means differed from each other
- But as noted before, if you test all combinations, you increase your chance of rejecting the null incorrectly
- To be conservative, we reduce the level of α , that is we will reject the p-value at a level smaller than the original α

Bonferroni method for multiple comparisons

- The Bonferroni method divides α by the number of possible pairs of tests

$$\alpha^* = \frac{\alpha}{\binom{k}{2}}$$

- Example: if you have 3 groups and you started with $\alpha=0.05$ then $\alpha^* = 0.05 / (3 \text{ choose } 2)$
 $= 0.05 / 3 = 0.01677$
- This means that you will only reject if $p < 0.017$

Multiple comparisons with ANOVA

- Use a t-test, but use the within group variance s_w^2 that weights over all the groups (not just the 2 being examined)
- The test statistic for each pair of means is:

$$t_{ij} = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{s_w^2 \left[(1/n_i) + (1/n_j) \right]}}$$

and the degrees of freedom are $n-k$ where n is the total number of observations and k is the total number of groups (note difference from regular t-test)

- Reject if the p-value is $< \alpha^*$
 - (Note: This is if you are doing the test by hand; if you use Stata option Bonferroni reject if $p < \alpha$)

Multiple comparisons

```
. oneway cd4count lastalc_3, bonferroni
```

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	344571.162	2	172285.581	2.45	0.0867
Within groups	69660550.3	991	70293.189		
Total	70005121.5	993	70498.6118		

Bartlett's test for equal variances: $\chi^2(2) = 2.4514$ Prob> $\chi^2 = 0.294$

Comparison of CD4Count by RECODE of lastalc (E1. Last time took alcohol)
(Bonferroni)

Row Mean - Col Mean	Never	>1 year
>1 year	-11.7697 1.000	
Within t	32.7188 0.239	44.4884 0.174

Difference between the 2 means

p-value for the difference, already adjusted
for the fact that you are doing multiple
comparisons (so reject if $p < \alpha$)

Statistical hypothesis tests

Data and comparison type	Alternative hypotheses	Test and Stata command
Numerical; One mean	$H_a: \mu \neq \mu_a$ (two-sided) $H_a: \mu > \mu_a$ or $\mu < \mu_a$ (one-sided)	Z or t-test • <code>ttest var1=hypoth val.*</code>
Numerical; Two means, <u>paired data</u>	$H_a: \mu_1 \neq \mu_2$ (two-sided) $H_a: \mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	Paired t-test • <code>ttest var1=var2*</code>
Numerical; Two means, independent data	$H_a: \mu_1 \neq \mu_2$ (two-sided) $H_a: \mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	T-test (equal or unequal variance) • <code>ttest var1, by(byvar) unequal</code>
Numerical, Two or more means, independent data	$H_a: \mu_1 \neq \mu_2$ or $\mu_1 \neq \mu_3$ or $\mu_2 \neq \mu_3$ etc.	ANOVA • <code>oneway var1 byvar</code>
Dichotomous; One proportion	$H_a: p \neq p_a$ (two-sided) $H_a: p > p_a$ or $p < p_a$ (one-sided)	Proportion test • <code>prtest var1=hypoth value*</code> • <code>bitest var1=hypoth value</code>
Dichotomous; two proportions	$H_a: p_1 \neq p_2$ (two-sided) $H_a: p_1 > p_2$ (one-sided)	Proportion test (z-test) • <code>prtest var1, by(byvar)</code>
Categorical by categorical (nxk)		

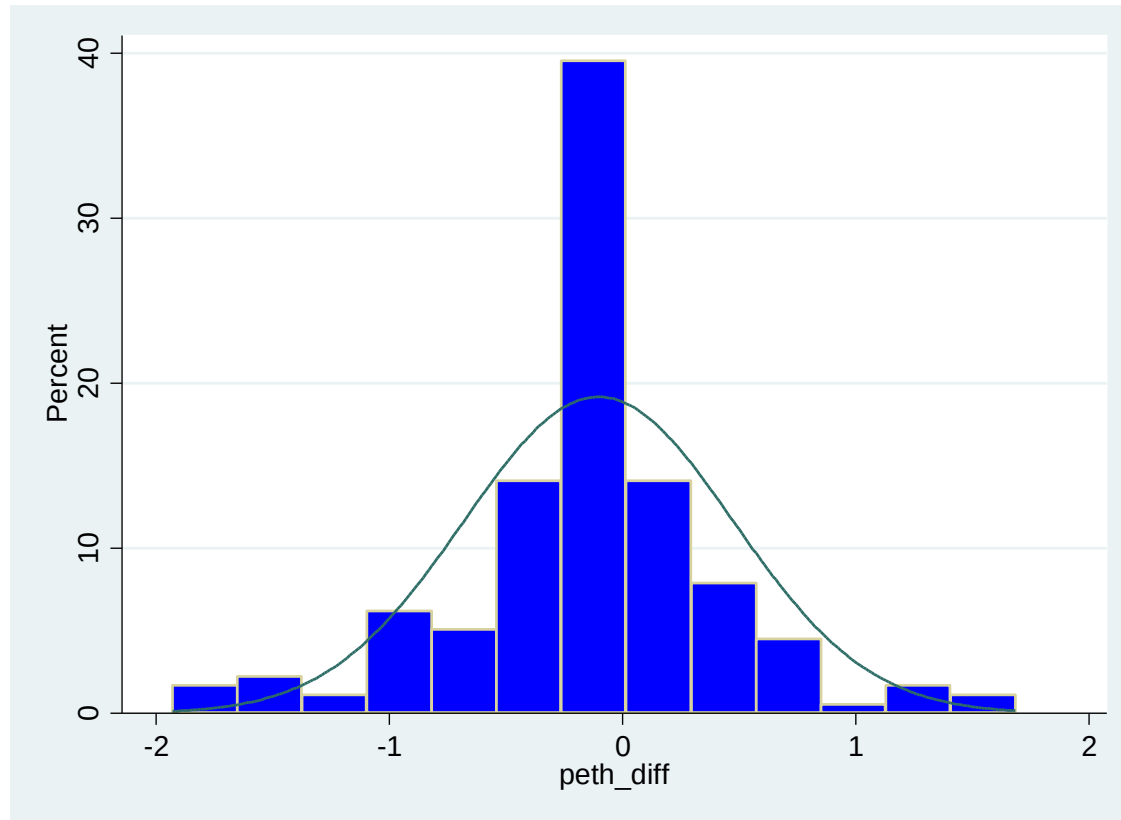
Parametric hypothesis test assumptions

- The hypothesis tests that use the z-statistic (i.e. when σ is known) assume that the underlying distribution of the parameter we are estimating (sample mean, sample proportion) is approximately normal.
 - True under the CLT if n is large enough.
- However, we usually do not know σ , and we use s^2 and compare our test statistic to the t-distribution. In theory, for this to work, the underlying distribution of the data must be normal, but in practicality, if n is fairly large and there are no extreme outliers, the t-test is valid.

Test assumptions

- However, if the data are not normally distributed, the t-test is not the most powerful test to use. (Note: less powerful does not mean invalid)
 - E.g. outliers will inflate the sample variance, decreasing the test statistic, thereby decreasing the chances of rejecting the null when it is false.
- Non-parametric tests do not rely on assuming a distribution for the data and therefore can help with this.
- Note that the independence of your observations is more critical than normality.
 - If your data points are not independent and you treat them as if they are, you will be acting like you have more data than you actually do (making you more likely to reject the null)

Differences log PEth (from 0-3 months)



* Using BREATH wide data.dta from Lecture 6 *
hist peth_diff, fcolor(blue) normal percent

Nonparametric tests for paired observations

- The Sign test
 - For paired or matched observations (analogous to the paired t-test)
 - $H_0 : \text{median}_1 = \text{median}_2$
 - Most useful when
 - the sample size is small
 - OR the distribution of differences is very skewed

Nonparametric tests for paired observations

- The Sign test
 - The differences between the pairs are given a sign:
 - + if a positive difference
 - if a negative difference
 - nothing if the difference=0
 - Count the number of +s , denoted by D

Nonparametric tests for paired observations

- Under H_0 , $\frac{1}{2}$ the differences will be $+s$ and $\frac{1}{2}$ will be $-s$
 - That is, $D/n = .5$
- This is equivalent to saying that each difference is a Bernoulli random variable, that is, each is $+$ or $-$ with probability $p=.5$
- Then the total number of $+s$ (D) is a binomial random variable with $p=0.5$ and with n trials

Nonparametric tests for paired observations

- So then the p-value for the hypothesis test is the probability of observing D + differences if the true distribution is binomial with parameters n and $p=0.5$
- $P(X=D)$ with n trials and $p=0.5$
- You could use the `binomialtail` function
- For a one-sided hypothesis:
 - `di binomialtail(n,D,.5)`
- For a two-sided hypothesis:
 - `di 2*binomialtail(n,D,.5)`

PEth results on 2 interviews

```
list studyid peth0_log10 peth3_log10 peth_diff in 1/10
```

	studyid	peth0~10	peth3~10	peth_diff	sign
1.	MBB2001	2.054766	1.950681	-.104085	-
2.	MBB2002	.30103	.30103	0	.
3.	MBB2006	.30103	.30103	0	.
4.	MBB2007	.30103	.30103	0	.
5.	MBB2008	1.388012	1.575477	.1874642	+
6.	MBB2011	.30103	.30103	0	.
7.	MBB2013	.30103	.30103	0	.
8.	MBB2016	2.122216	.30103	-1.821186	-
9.	MBB2021	1.341533	1.348207	.0066741	+
10.	MBB2022	.30103	1.111766	.8107364	+

```
** Using BREATH wide data.dta
```

```
** 1st 10 observations *
```

Sign test

```
. gen spd=sign(peth_diff)
. tab spd
```

spd	Freq.	Percent	Cum.
-1	76	42.94	42.94
0	45	25.42	68.36
1	56	31.64	100.00
Total	177	100.00	

- D=76 negative differences
- N=177-45=132 (don't count the 45 ties)
- Using binomial distribution

```
. . di 2*binomialtail(132,76,.5)
.09781221
```




In Stata

NOTE that there is only 1 = in the command!

signtest var1=var2

```
. signtest peth3_log10=peth0_log10
```

Sign test

sign	observed	expected
-----+-----		
positive	56	66
negative	76	66
zero	45	45
-----+-----		
all	177	177

One-sided tests:

Ho: median of peth3_l~10 - peth0_log10 = 0 vs.

Ha: median of peth3_l~10 - peth0_log10 > 0

Pr(#positive >= 56) =

Binomial(n = 132, x >= 56, p = 0.5) = 0.9664

Ho: median of peth3_l~10 - peth0_log10 = 0 vs.

Ha: median of peth3_l~10 - peth0_log10 < 0

Pr(#negative >= 76) =

Binomial(n = 132, x >= 76, p = 0.5) = 0.0489

Two-sided test:

Ho: median of peth3_l~10 - peth0_log10 = 0 vs.

Ha: median of peth3_l~10 - peth0_log10 != 0

Pr(#positive >= 76 or #negative >= 76) =

min(1, 2*Binomial(n = 132, x >= 76, p = 0.5)) = 0.0978

Uses the larger of the number of positive or negative signed pairs

Nonparametric tests for paired observations

- Note that the Sign test can be used for ordinal data
- The sign test does not account for the magnitude of the difference in the outcome variable
- Another test, the Wilcoxon Signed-Rank Test, ranks the differences in the pairs

Nonparametric tests for paired observations: Wilcoxon Signed-Rank Test

- Null hypothesis :
$$\text{median}_1 = \text{median}_2$$
- The differences in the pairs are ranked
- Ties are given the average rank of the tied observations
- Each rank is assigned a sign (+/-) depending on whether the difference is positive or negative
- The ranks for the positives and the negatives are added up separately; the absolute value of the smaller sum of the ranks is called T

Nonparametric tests for paired observations

- T follows a normal distribution with mean m_T and standard deviation σ_T

$m_T = n*(n+1)/4$ (the rank sum if both medians were equal)

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

The test statistic $z_T = (T - m_T) / \sigma_T$

Compare to the standard normal distribution

For $n < 12$, use the exact distribution, table A.6

```
egen pethrank_diff=rank(peth_diff)
```

```
. list studyid peth3_log10 peth0_log10 peth_diff pethrank_diff in  
1/20
```

	studyid	peth3~10	peth0~10	peth_diff	pethra~f
1.	MBB2001	1.950681	2.054766	-.104085	65
2.	MBB2002	.30103	.30103	0	99
3.	MBB2006	.30103	.30103	0	99
4.	MBB2007	.30103	.30103	0	99
5.	MBB2008	1.575477	1.388012	.1874642	139
6.	MBB2011	.30103	.30103	0	99
7.	MBB2013	.30103	.30103	0	99
8.	MBB2016	.30103	2.122216	-1.821186	3
9.	MBB2021	1.348207	1.341533	.0066741	124
10.	MBB2022	1.111766	.30103	.8107364	170
11.	MBB2025	1.026124	2.054287	-1.028163	11
12.	MBB2026	1.717338	2.329398	-.6120603	26
13.	MBB2027	1.659488	.30103	1.358458	174
14.	MBB2029	.30103	.30103	0	99
15.	MBB2030	.30103	.30103	0	99
16.	MBB2032	.30103	.30103	0	99
17.	MBB2035	.30103	1.245883	-.9448526	16
18.	MBB2036	.30103	.30103	0	99
19.	MBB2040	1.650939	2.472025	-.821086	20
20.	MBB2041	2.856124	2.512551	.3435733	152

```
signrank var1 = var2
```

```
. signrank peth3_log10=peth0_log10
```

Wilcoxon signed-rank test

sign	obs	sum ranks	expected
-----+-----			
positive	56	5882	7359
negative	76	8836	7359
zero	45	1035	1035
-----+-----			
all	177	15753	15753

unadjusted variance	466026.25
adjustment for ties	-0.13
adjustment for zeros	-7848.75

adjusted variance	458177.38

Ho: peth3_log10 = peth0_log10

z = -2.182

Prob > |z| = 0.0291

This is a two-sided p-value
arrived at using

```
di 2*(normal(-2.182))  
.02910953
```

If you wanted a one-sided
test, use

```
. di normal(-2.182)  
.01455477
```

Nonparametric tests for two independent samples (e.g. 2 arms of the study)

- The Wilcoxon Rank Sum Test
 - Also called the Mann-Whitney U test
- Null hypothesis :
$$\text{median}_1 = \text{median}_2$$
- Samples from independent populations – analogous to the t-test
- Assumes that the distributions of the 2 groups have the same shape

Nonparametric tests for two independent samples

- The entire sample (including the members of both groups) is ranked
- Average rank is given to ties
- Sum the ranks for each of the 2 samples – smaller sum is W
- The test statistic $z_W = (W - m_W) / \sigma_W$ is compared to the normal distribution (see P+G page 310 for the formula)
- If the sample sizes are small (<10), exact distributions are needed – Table A.7



ranksum var, by(byvar)

```
. ranksum peth_log10, by(studyarm)
```

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

studyarm	obs	rank sum	expected
Cohort asses	181	25643.5	26064
Min assessed	106	15684.5	15264
combined	287	41328	41328

unadjusted variance	460464.00
adjustment for ties	-20670.38
adjusted variance	439793.62

This is a two-sided p-value arrived at using

```
di 2*normal(-0.634)
.52608082
```

If you wanted a one-sided test, use

```
.di normal(-0.634)
.26304041
```

Ho: peth_~10(studyarm==Cohort assessed) = peth_~10(studyarm==Min assessed)

z = -0.634

Prob > |z| = 0.5260

** Using BREATH cohort and comparisons_v2.dta **

Nonparametric tests for multiple independent samples

- The Kruskal-Wallis test extends the Wilcoxon rank sum test to 2 or more independent samples
 - You could use the Kruskal-Wallis with 2 independent samples and reach the same conclusion as if you had used the Wilcoxon
- Analogous to one-way analysis of variance

Nonparametric tests for independent samples (Kruskal Wallis)

```
kwallis      var , by(byvar)
```

```
. kwallis cd4count, by(lastalc_3)
```

Kruskal-Wallis equality-of-populations rank test

lastalc_3	Obs	Rank Sum
Never	373	181395.00
>1 year ago	180	83338.00
Within the past year	441	229782.00

```
chi-squared =      6.134 with 2 d.f.  
probability =      0.0466
```

```
chi-squared with ties =      6.134 with 2 d.f.  
probability =      0.0466
```

Parametric vs. non-parametric (distribution free) tests

- Non parametric tests:
 - No normality requirement
 - Do require that the underlying distributions being compared have the same basic shape
 - Ranks are less sensitive to outliers and to measurement error
- If the underlying distributions are approximately normal, then the parametric tests are more powerful

Statistical hypothesis tests

Data and comparison type	Alternative hypotheses	Parametric test Stata command	Non-parametric test Stata command
Numerical; One mean	$H_a: \mu \neq \mu_a$ (two-sided) $H_a: \mu > \mu_a$ or $\mu < \mu_a$ (one-sided)	Z or t-test • <code>ttest var1=hypoth val.*</code>	
Numerical; Two means, <u>paired data</u>	$H_a: \mu_1 \neq \mu_2$ (two-sided) $H_a: \mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	Paired t-test • <code>ttest var1=var2*</code>	Sign test • <code>signtest var1=var2</code> • Wilcoxon Signed-Rank <code>signrank var1=var2</code>)
Numerical; Two means, independent data	$H_a: \mu_1 \neq \mu_2$ (two-sided) $H_a: \mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	T-test (equal or unequal variance) • <code>ttest var1, by(byvar)</code> <code>unequal</code>	Wilcoxon rank-sum test • <code>ranksum var1, by(byvar)</code>
Numerical, Two or more means, independent data	$H_a: \mu_1 \neq \mu_2$ or $\mu_1 \neq \mu_3$ or $\mu_2 \neq \mu_3$ etc.	ANOVA • <code>oneway var1 byvar</code>	Kruskal Wallis test • <code>kwallis var1, by(byvar)</code>
Dichotomous; One proportion	$H_a: p \neq p_a$ (two-sided) $H_a: p > p_a$ or $p < p_a$ (one-sided)	Proportion test • <code>prtest var1=hypoth value*</code> • <code>bitest var1=hypoth value</code>	
Dichotomous; two proportions	$H_a: p_1 \neq p_2$ (two-sided) $H_a: p_1 > p_2$ (one-sided)	Proportion test (z-test) • <code>prtest var1, by(byvar)</code>	
Categorical by categorical (nxk)	H_a : The rows not independent of the columns		

For next time

- Read Pagano and Gauvreau
 - Pagano and Gauvreau Chapters 12-13 (review)
 - Pagano and Gauvreau Chapter 15