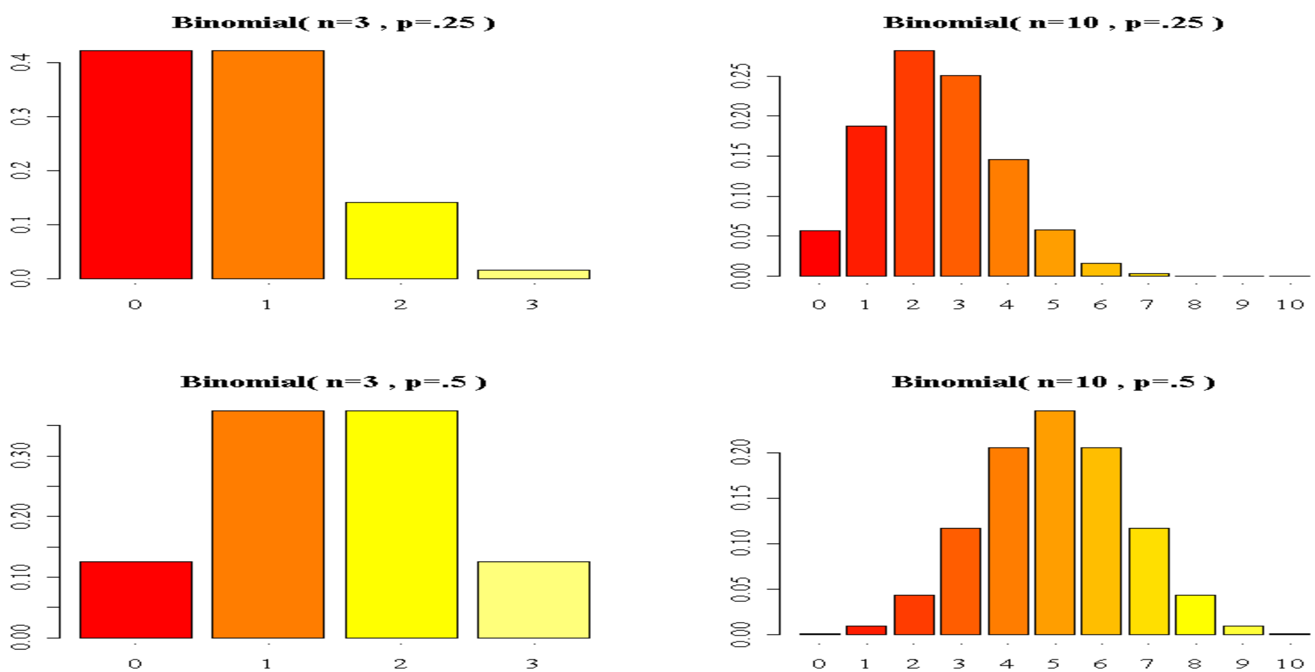
PBIO 504

Lecture 3: T-Tests

Textbook Chapters 9.3 and 11

$$\mu = np \qquad \sigma = \sqrt{np(1-p)}$$

- The Central Limit Theorem says that random variables that are the averages of large number of independent identically distributed random variables are approximately normally distributed. Based on this *we can approximate the binomial distribution by a normal distribution with the mean and standard deviation (see above), when n is large.*



Hypothesis Testing

Population: B cell lymphoma

(null) H_o: New trt = Control trt

(alternative) H_a: New trt $\neq$ Control trt

Statistical significance level: $\boldsymbol{\alpha} = 0.05$

Power: $\mathbf{1–\boldsymbol{\beta}} = 0.80$

Hypothesis Testing

α = probability of making Type I error

= statistical significance level

= probability of deciding that H_a: New $\neq$ C (the alternative hypothesis) is true, when in fact the null hypothesis is true H_o: New = C

Hypothesis Testing

β = probability of making Type II error

β = probability of deciding that H_o is true (New = C), when in fact H_a is true (New $\neq$ C)

Power = 1**– β**

Power is the probability of deciding that H_a is true when in fact it is true.

Hypothesis Testing

- What does **α =** 0.05 really mean?

If H_o: New = C is true, then the statistical test will falsely conclude that the alternative H_a: New $\neq$ C is true, 5% of the time.

In other words, if we repeat the test many times with different samples from the same population, the probability of type I error is 0.05

Hypothesis Testing
Specific Alternative: Power

- The response rate is hypothesized to be more than 40% if the new therapy is deemed effective

- Power 80% means that 80% of these tests will conclude that the H_a: New $\neq$ C is true, actually further, that New trt > C is true because the new therapy is effective as described above.

What is a Test Statistic?

- *Test Statistic* is a statistic (formula) that helps us to perform tests.

$$T = \frac{\text{Estimated treatment difference based on data}}{\text{The standard deviation of the estimate}}$$

After data is colleted, the calculated T value is compared with a critical level based on the distribution of test statistic T and the given type I and II errors

Sampling Distribution

- Recall that

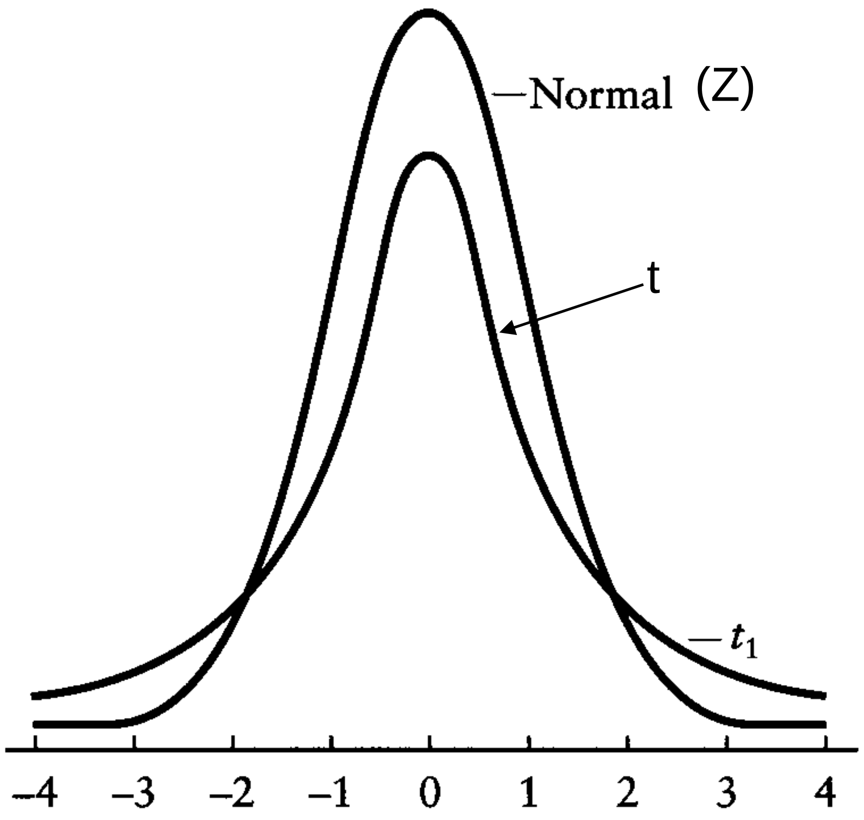
  $Z = (\bar{X} - \mu) / (\sigma / \sqrt{n})$ where σ is the population standard deviation
- If we don't know σ, then we can use the observed standard deviation from the sample population, notated "s".
- This yields the Student t distribution*,

  $T = (\bar{X} - \mu) / (s / \sqrt{n})$

* Named after Gossett who published this under the pseudonym: Student

Compared to Z, the t distribution has:

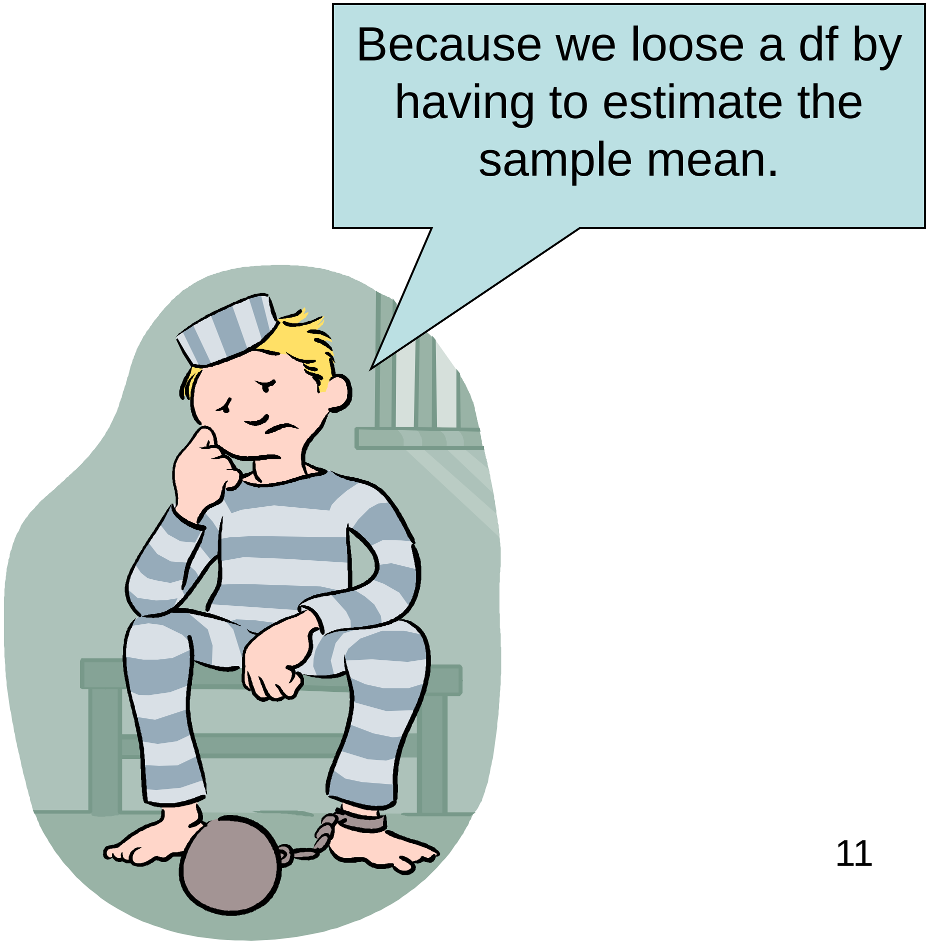
- a mode and symmetrical shape, like Z;
- area under the curve equals 1, like Z; but
- somewhat "thicker" tails than Z, reflecting extra variability that comes from ***s*** as an estimate of ***σ***.



Degrees of Freedom (df)

The t Distribution has Degrees of Freedom

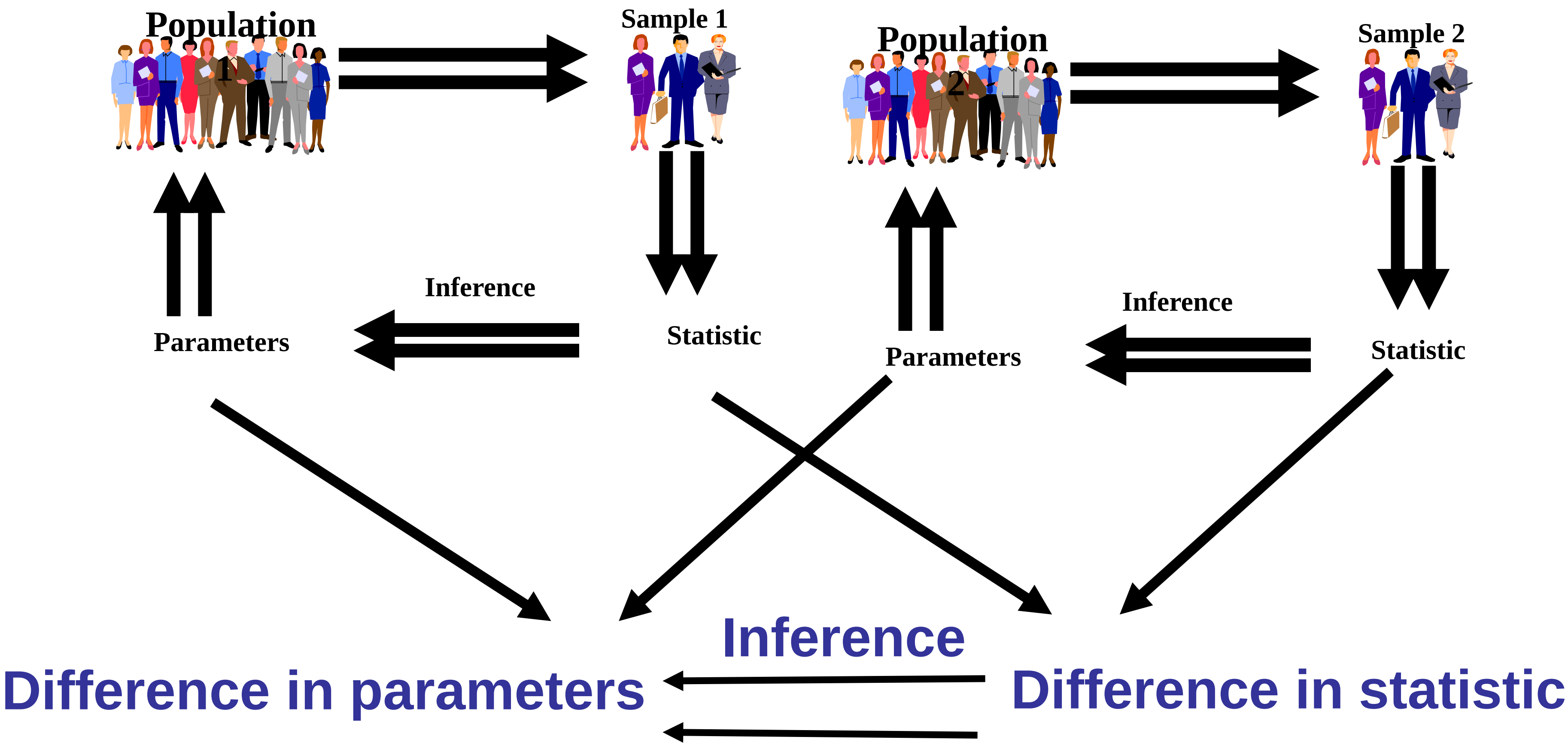
- Degrees of freedom measure the amount of information that is available to estimate ***σ,*** i.e. the reliability of ***s*** as an estimate.
- For the t distribution, ***df = n-1***
- Why is it ***n-1*** and not ***n***?



Degrees of Freedom (cont.)

- Each value of ***df*** has a different t distribution.
- Small ***df*** distributions are more spread out than larger ones.
- As the ***df*** increase, the t distribution approaches the Z distribution, since ***s*** becomes a more reliable estimate of $\boldsymbol{\sigma}$.

t Test for Comparing 2 Means



t-Test for Comparing 2 Means

- When we want to compare means (μ_1, μ_2) from two different study populations, we use the Student t-test (also called the 2 sample t test).
- The test determines whether the difference between the 2 sample means ($\bar{X}_1 - \bar{X}_2$) is too large to have occurred by chance.
- The Null Hypothesis for this test is that $\mu_1 = \mu_2$. The Alternative Hypothesis is that μ_1 and μ_2 are not equal $\mu_1 \neq \mu_2$

Null and Alternative Hypotheses

- Null hypothesis for t-test

$$H_0: \mu_1 = \mu_2.$$

- Alternative hypothesis for t-test

$$H_A: \mu_1 \neq \mu_2.$$

Accept or Reject Based on p-value

- We will reject the null hypothesis if the p-value from the t test is less than the pre-set level of significance (alpha, or ***α***).
- ***α*** is the level of chance (probability) that we are willing to accept for wrongly rejecting the null.
- ***α*** is often set at 0.05, but it does not have to be this value. It can be set at whatever value makes sense for the specific study we are conducting.
- If $p \le \alpha$, then reject H_0 and conclude H_A.
- If $p > \alpha$, then we fail to reject H_0, (or some speakers may say that they accept H_0).

Levels of Alpha

- When we say ***α*** is 0.05, we mean that we are willing to accept a 5% chance of wrongly rejecting the null hypothesis.

- In research, 0.05 might not always be the best value, other values for type I error may be more appropriate.

Comparison of Two Populations Means

- For this 2-sample t test, there are assumptions that must be met: data points are independent and normally distributed.
- Recall $H_0\!: \mu_1 = \mu_2$.

$$T = (\bar{X}_1 - \bar{X}_2) \,/\, \sqrt{\{s^2_p\,[\,(1/n_1) + (1/n_2)\,]\,\}}$$

where s^2_p is the pooled variance of the two samples.

Note that T depends on the two sample means, the two sample sizes, and the sample variance (s^2), where $s = \sqrt{s^2}$

Equal and Unequal Variances

- The formula just discussed for T assumes that the variances of the two populations are equal, and this fact is used when pooling the variance estimates across the samples.
- If this is not true, then a special formula has to be used for T, as follows:
- $T = (\bar{X}_1 - \bar{X}_2) / \sqrt{(s^2_1 / n_1) + (s^2_2 / n_2)}$
- STATA allows you to select either formula, in case your study does or does not conform to the equal variance assumption.

t-Test Example

- In a study of 207 male and 127 female diabetics, one of the factors being investigated was body mass index (BMI), which is weight (kg) / height (m^2). Compare the mean BMI for males vs. females.
- Male sample mean BMI was 26.4 and std 5.3
- Female sample mean BMI was 25.4 and std 5.2
- Notice that the stds estimates are approx equal
- T=2.87, df =(207+127)-2=332, p=0.0022 < 0.05 and we reject the null hypothesis and conclude that the mean BMI for M and F are different

Another Example

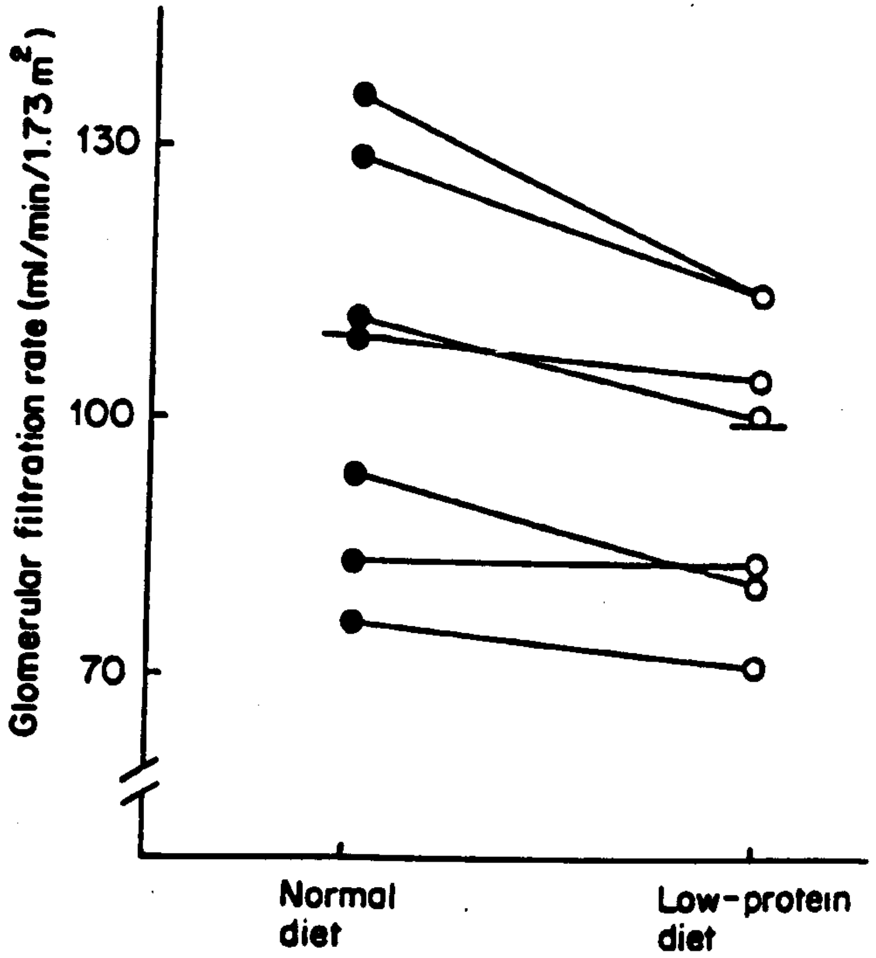
- In a study of meningitis in two populations of patients, with HIV infection (group 1) and without HIV (group 2), investigators were interested in learning if the mean ages were the same or different.
- Since the two groups are independent, and age was normally distributed with equal variance in the populations, the two-sample t-test is appropriate to test H_0: the age means are equal.

The Paired t Test

- This is a special case where the t distribution is used for the situation of “paired” samples.
- This means that every subject observed in the first group is observed again in the second group of data points.
- *<u>Example</u>: a person's blood pressure is measured at baseline, and again after an intervention.*
- Pairing is a useful design feature because it can reduce variability in a study, since each person serves as his or her own control subject.
- First we calculate the differences btw the paired observations and then we calc the average.

Example from a Previous Lecture

Glomerular filtration rate during normal and low protein diets in diabetic patients



- Coehn et al. (1987) showed the changes in kidney function of patients who were first fed a normal diet and then switched to a low protein diet.
- A line connects the pair of data points for each of the 7 subjects in the study.
- The statistical problem is this: ***is the mean change in filtration rate significant?***

The Paired T Statistic

- In this type of study, the measure of interest is the difference in the means between the two time points.
- H_0: the mean of the differences is zero
- $T = \bar{d} / (s_d / \sqrt{n})$

  where n is the number of pairs and s_d is the standard deviation of $\bar{d}$
- df= n-1

Paired t Test Example

- A group of 62 men who had survived a previous heart attack were tested on a treadmill, to see if a deep breathing exercise prior to walking delayed the onset of angina (heart pains).
- Each man was measured on the treadmill before and after the deep breathing exercise.

Paired t Test Solution

- The mean time to angina, before the deep breathing, was 93 seconds (std=223) on the treadmill.
- The mean time to angina, after the deep breathing, was 587 seconds (std=180) on the treadmill.
- diff = 587-93 = 494 seconds
- T=2.05, df=61, $p=0.045 < 0.05$
- ***How do we interpret this result?***

  There is a stat sign diff btw mean time to angina between the two time points.

T test with STATA

Type help(ttest) into STATA for more details

Syntax

One-sample t test

```
ttest varname == # [if] [in] [, level(#)]
```

Two-sample t test using groups

```
ttest varname [if] [in] , by(groupvar) [options1]
```

Two-sample t test using variables

```
ttest varname1 == varname2 [if] [in], unpaired [unequal welch level(#)]
```

Paired t test

```
ttest varname1 == varname2 [if] [in] [, level(#)]
```