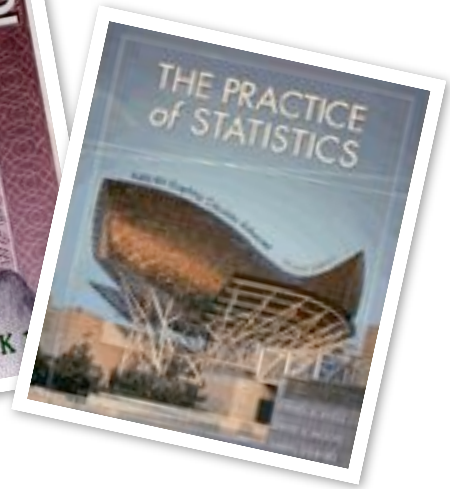
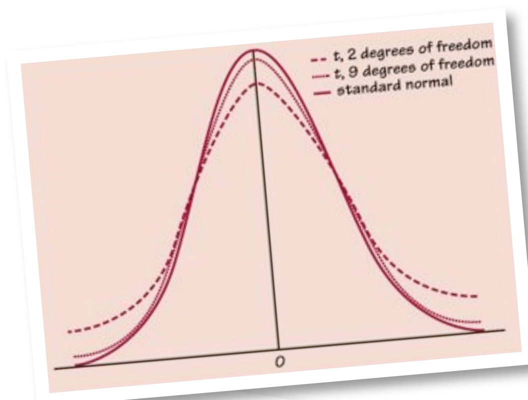
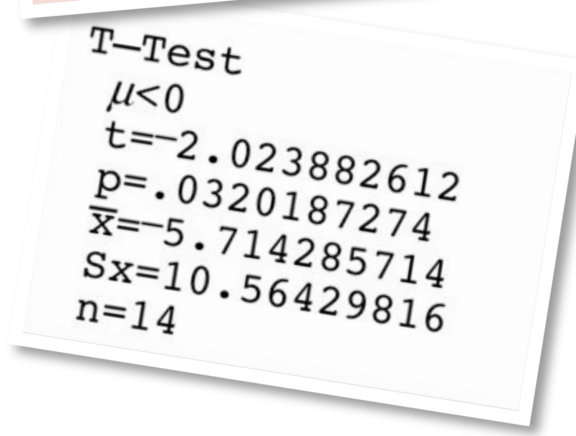


Chapter 11

Inference for Means

In Chapter 10, we learned the logic behind inferential procedures. In this chapter, we will apply that logic to inference involving means. We will learn how to build confidence intervals and perform significance tests for one mean as well as comparisons between two means. Further, we will be introduced to a new distribution that we can use when we do not know the population standard deviation.

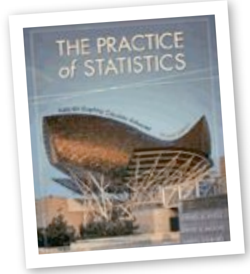




Inference For Means:

- 11.1 The t - Distribution
- 11.1 Inference for a Mean
- 11.2 Comparing Two Means

AP Statistics Chapter 11: Inference for Means



"A statistical analysis, properly conducted, is a delicate dissection of uncertainties, a surgery of suppositions."
~M.J. Moroney

Tentative Lesson Guide					
Date		**Stats**	**Lesson**	**Assignment**	**Done**
Wed	2/14	11.1	t Distributions	Rd 616-619 Do 1-5	
Thu	2/15	11.1	t intervals and tests	Rd 621-628 Do 7-11	
Fri	2/16	11.1	Practice	Practice Problems	
Mon	2/19		***No School***		
Tues	2/20	**Qz**	**Quiz 11.1**	Rd 648-656 Do 37-38	
Wed	2/21	11.1	Matched Pairs t test	Rd 628-640 Do 12-17	
Thu	2/22	11.2	Comparing Two Means	Rd 658-667 Do 39-43,47, 49	
Fri	2/23	**Qz**	**Quiz 11.2**	Rd 667-668 Do 50, 53, 55	
Mon	2/26	Rev	Review Ch 11	Rd 673-674 Do 62-65, 72	
Tues	2/27	**Ex**	**Exam Chapter 11**	Online Quiz Due	

Note:
The purpose of this guide is to help you organize your studies for this chapter.
The schedule and assignments may change slightly.

Keep your homework organized and refer to this when you turn in your assignments at the end of the chapter.

Class Website:
Be sure to log on to the class website for notes, worksheets, links to our text companion site, etc.

http://web.mac.com/statsmonkey

Don't forget to take your online quiz!. Be sure to enter my email address correctly!
http://bcs.whfreeman.com/yates2e

My email address is:

jmmolesky@isd194.k12.mn.us

Chapter 11 Objectives and Skills:

These are the expectations for this chapter. You should be able to answer these questions and perform these tasks accurately and thoroughly. Although this is not an exhaustive review sheet, it gives a good idea of the "big picture" skills that you should have after completing this chapter. The more thoroughly and accurately you can complete these tasks, the better your preparation.

t-Distributions

- [ ] Describe the sampling distribution of x-bar when the population standard deviation is unknown.
- [ ] Describe t-distributions for different degrees of freedom. Note that the t-distribution becomes approximately normal as n approaches infinity.
- [ ] Find t-statistics and p-values for sample means.

Inference for a Single Mean

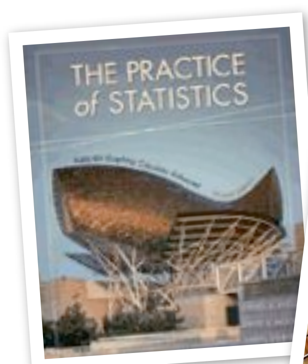
- [ ] Construct and interpret a level C confidence interval for a single mean when the population standard deviation is not known.
- [ ] Conduct a significance test for a claim about a single mean.
- [ ] Conduct a matched pairs t-test for the mean difference in a matched pairs setting.

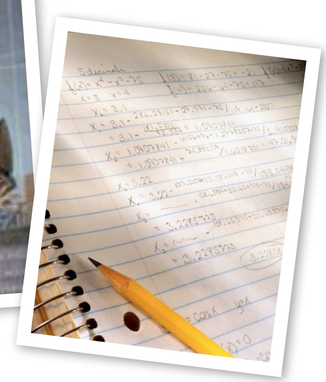
Inference for Two Means

- [ ] Describe the sampling distribution for the difference between sample means from two independent populations.
- [ ] Calculate and interpret a Level C confidence interval for the difference between two means.
- [ ] Conduct a two-sample t-test for the difference between two means.

Calculator Procedures

- [ ] Be able to calculate and interpret Confidence Intervals for means using your graphing calculator.
- [ ] Be able to perform a one- or two-sample t-test using your graphing calculator.
- [ ] Recognize that the graphing calculator is only a tool to provide you with the test statistic and p-value...it is up to YOU to interpret the results!





11.1: Introduction to the t-Distributions

In Chapter 10, we performed inferential procedures on the population mean under the assumption that we knew the population standard deviation. However, in practice, the population standard deviation is not known--if it was, that would imply you knew the population mean and the inference would be somewhat unnecessary!

In this chapter, we'll perform inferential procedures on the mean without knowing the population standard deviation. However, to do so means we need to learn a new type of distribution. Since we don't know the population standard deviation, we can no longer rely on the normal distribution for our calculations. Thankfully, William Sealy Gosset solved that problem for us...

William Sealy Gosset - The Student's t-Distribution

What would cause the head brewer of the Guinness brewery in Dublin, Ireland, to not only use statistics but also to invent new statistical methods? Why...the search for better beer, of course!

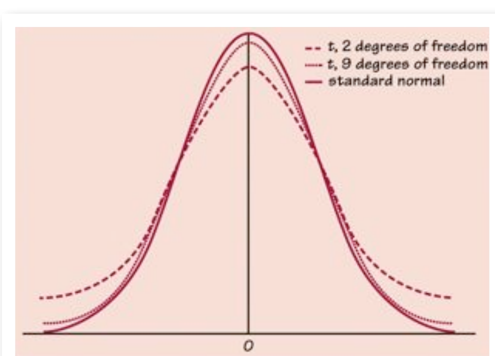
William S. Gosset (1876-1937), fresh from Oxford University, joined Guinness as a brewer in 1899. He soon became involved in experiments and in statistics to understand the data from those experiments. What are the best varieties of barley and hops for brewing? How should they be grown, dried, and stored? The results of field experiments to answer these questions varied. Gosset faced the problem we noted in using the z-test--he didn't know the population standard deviation. . Further, he noticed that replacing sigma by the sample standard deviation, s, in the z-score formula and calling the result roughly Normal wasn't good enough.

After much work, Gosset developed what we now call the *t*-distributions. His new *t* test identified the best barley variety, and Guinness promptly bought up all the available seed. Guinness allowed Gosset to publish his statistical discoveries, but not under his own name. Therefore, Gosset used the pseudonym "Student" on his work, leading to the name "Student's *t*-Distributions." {Source: YMS 3e, Chapter 12}

Student's *t*-Distribution

In probability and statistics, the t-distribution (or Student's *t*-distribution) is a probability distribution that arises in the problem of estimating the mean of a normally distributed population when the sample size is small. Student's distribution arises when the population standard deviation is unknown and has to be estimated from the data--as is the case in nearly all practical statistical work.

If the population standard deviation is unknown, we must estimate it using the sample standard deviation, *s*. The value of *s* may not be close to sigma--especially if *n* is small. Further, *s* may vary from sample to sample. As a result, the use of s in place of sigma introduces extra variability into our problem. Due to the extra variability, the *t*-distribution is more spread out than the normal (z) distribution.



11.1: *t*-Distributions

Just like normal distributions, there are many different *t*-distributions. While normal distributions are distinguished by their mean and standard deviation, *t*-distributions are distinguished by a value called the *degrees of freedom (df)*. This value is determined by the setting in which the *t*-distribution is applied.

Properties of *t*-Distributions
☑
☑
☑
☑

When dealing with t-distributions, it is important to note the degrees of freedom that correspond to your particular problem. The df define which t-distribution you are dealing with and determine the critical values and p-values for the problem. The t-table, like the z-table, can be used to determine p-values and critical scores in statistical problems. However, you have to read the t-table a little bit differently...

Table B *t* **distribution critical values**

	Tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											

11.1: Estimating a Mean

Inference Toolbox: *Confidence Interval for a Single Mean μ*

To Construct a Confidence Interval for a Single Mean μ:

- ☑
- ☑
- ☑
- ☑

Example:

The article "*Increased Vital and Total Lung Capacity in Tibetan Compared to Han Residents of Lhasa*" compared various physical characteristics of people living at high altitudes to those living at sea level. According to the statistics in the article, a sample of 38 Tibetan residents resulted in a mean lung capacity of 6.8 liters with s=1.17 liters. 43 Han residents had an average lung capacity of 6.24 liters with s=1.18 liters.

Calculate and interpret a 95% confidence interval for the lung capacity of Tibetan residents.

☑ Parameter of Interest:
☑ Conditions:
☑ 95% Confidence Interval for μ {df=______ t*=_______}

☑ Interpret:

Calculate and interpret a 90% confidence interval for the lung capacity of Han residents.

☑ Parameter of Interest:
☑ Conditions:
☑ 95% Confidence Interval for μ {df=______ t*=_______}

☑ Interpret:

11.1: Tests of Significance for a Single Mean

Inference Toolbox: ***Test of Significance for a Single Mean μ***

To Test a Claim about an unknown Population Mean μ:

- ☑
- ☑
- ☑
- ☑
- ☑
- ☑

Example...Sweet...

Diet colas use artificial sweeteners to avoid sugar. These sweeteners may lose their sweetness over time. Therefore, manufacturers test new colas for loss of sweetness before marketing them. Trained testers sip the colas and rank them on a "sweetness scale" of 1 to 10. The cola is then stored for a month at high temperatures to simulate the effect of four month's storage at room temperature. Each taster scores the cola again after storage and the difference is noted (before storage - after storage). Positive differences indicate a loss in sweetness--the bigger the difference, the bigger the loss.

Here are the sweetness losses for a new cola, as measured by 10 trained testers:

2.0	**0.4**	**0.7**	**2.0**	**-0.4**	**2.2**	**-1.3**	**1.2**	**1.1**	**2.3**

Do these data provide evidence that the cola lost sweetness in storage? Assume 5% significance.

11.1: Paired *t* Tests

Experiments can be designed to yield paired data in a number of different ways. Some studies involve using the same group of individuals with measurements recorded before and after some intervening treatment. Others use naturally occurring pairs, such as twins or husband-wife, and some construct pairs by matching on factors with effects that might otherwise obscure differences between the two populations of interest. Paired samples often provide more information than would independent samples, because extraneous effects are screened out. Therefore, comparative matched-pairs studies can be a powerful statistical tool with inferential results that are often more convincing than single sample results.

To compare the responses to two treatments in a matched pairs design, we simply apply the one-sample t-procedures to the observed differences between the sample pairs.

<u>Example</u>: Ahhhh! I need more coffee!

To determine whether or not caffeine dependence was a real phenomenon, E, Strain ("Caffeine dependence syndrome: evidence from case histories and experimental evaluation," J. of Am. Medical Assoc., 272(1994), pp.1604-1607) conducted a study in which 11 subjects were asked to perform a task twice - once under the influence of caffeine, and once under the influence of a placebo. The study measured how fast the subjects could repeatedly push a button when under the effects of the two treatments. The data follow...Button data is given in beats per minute the subject achieved. Do these data provide significant evidence that caffeine had a positive effect on the number of beats per minute?

Beats Caffeine	281	284	300	421	240	294	377	345	303	340	408
Beats Placebo	201	262	283	290	259	291	354	346	283	391	411
Difference											

11.1: Practice

1. The usual chirp rate for male field crickets varies around a mean of 60 chirps/sec. To investigate whether chirp rate was related to nutritional status, investigators fed 32 male crickets a high protein diet for 8 days, after which the chirp rate was measured. The mean rate for crickets on the diet was 109 chirps/sec with s=40. Is there evidence to suggest a high protein diet results in a higher chirp rate? Test at .01 significance level. After testing, construct and interpret a 99% confidence interval for the true chirp rate for crickets on a high protein diet.

2. Many consumers pay attention to stated nutritional contents on packaged foods when making purchases. Therefore, it is important that the information on the packages be accurate. A random sample of 12 frozen dinners was selected and calorie contents of each one was determined. The box reported the calorie content was 240. Do the following data provide evidence to suggest the calorie content may be higher?

255	244	239	242	265	245	259	248	225	226	251	233

3. The article "Caffeine Knowledge, Attitudes, and Consumption in Adult Women" (J. of Nutritional Ed. [1992]:179-184) reported the following summary statistics on daily caffeine consumption for a random sample of 47 adult women. .

Construct a 90% confidence interval for the true average caffeine consumption for adult women. The average was previously assumed to be 200mg. Does your interval support this belief?

4. Can taking chess lessons and playing chess daily improve memory? The article "The USA Junior Chess Olympics Research: Developing Memory and Verbal Reasoning" described a study in which 6th grade students who had not previously played chess took lessons and played daily for 9 months. Each student to a memory test before and after the program. Data from these tests follows. Do the data suggest chess may improve memory? Test using a matched-pairs t-test at the 5% significance level. Note: higher scores indicate better recall.

Pre	510	610	640	675	600	550	610	625	450	720	575	675
Post	850	790	850	775	700	775	700	850	690	775	540	680
Diff												

5. Much concern has been expressed in recent years regarding the practice of using nitrates as meat preservatives. In one study involving possible side effects of these chemicals, bacteria cultures were grown in a medium containing nitrates. The rate of uptake of radio-labeled amino acid was determined for each culture, yielding the following data:

7251	6871	9632	6866	9094	5849	8957	7978
7064	7494	7883	8178	7523	8724	7468	

The true average uptake for cultures without nitrates is 8000. Do the data suggest that the addition of nitrates results in a decrease in the true average uptake? Test at .10 significance.

11.2: Inference for Two Means

While the matched pairs pairs procedure allows us to perform inference on two sets of data that are the result of a matched setting, we will often encounter situations that require us to compare means that come from two independent populations. For example, how does the average ACT math score for males compare to the average ACT math score for females? How does the blood pressure for an experimental group taking a new medication compare to that of a placebo group? Questions like this can be answered through the use of two-sample procedures.

Notation:

	Parameters			**Statistics**	
Population	**Mean**	**Standard Deviation**	**Sample Size**	**Mean**	**Standard Deviation**
1					
2					

Like the one-sample procedures, two-sample procedures are based on the **Sampling Distribution** of the parameter of interest. In these cases, we are interested in the difference of the means of the two populations.

Sampling Distribution of $\bar{x}_1 - \bar{x}_2$

Inference Toolbox: ***Confidence Interval for a Difference between Two Means***

To Construct a Confidence Interval for a Difference between Means µ1-µ2:

☑

☑

☑

☑

Inference Toolbox: ***Two Sample Test for a Difference between Two Means***

Two Sample *t* Statistic

☑

11.2: Practice

1.Do children diagnosed with attention deficit/hyperactivity disorder have smaller brain volumes than children without the condition? The following data comes from research described in the article "Developmental Trajectories of Brain Volume Abnormalities in Children and Adolescents with ADHD". Brain scans were completed for 152 children with ADHD and 139 children of similar age without ADHD. Do these data provide convincing evidence that the mean brain volume of children with ADHD is smaller than that of children without ADHD?

	n	x-bar	s
With ADHD	152	1059.4	117.5
Without AHDH	139	1104.5	111.3

2. Does blood pressure tend to be higher in a doctor's office than when measured in a less stressful environment? The article "The Talking Effect and 'White Coat' Effect in Hypertensive Patients: Physical Effort or Emotional Content" described a study in which patients with high blood pressure were randomly assigned to one of two groups. Those in the first group (talking) were asked questions about their medical history and sources of stress in their lives minutes before their bp was measured. Those in the second group (counting) were asked to count aloud from 1 to 100 four times before their bp was measured. The following data for diastolic blood pressure appear in the paper:

Talking: $n = 8$ $\bar{x} = 108.75$ $s = 4.74$

Counting: $n = 8$ $\bar{x} = 102.25$ $s = 5.39$

Construct and interpret a 95% confidence interval for the difference in mean blood pressure.

3.The article "Does Smoking Cessation Lead to Weight Gain?" described an experiment in which 322 subjects, selected at random from those show successfully participated in a program to quit smoking were weighed at the beginning of the program and again one year later. The mean change in weight was 5.15 lb and the standard deviation was 11.45 lb. Is there sufficient evidence at =.05 to conclude that the true mean change is positive?

4.A researcher at the Medical College of Virginia conducted a study of 60 randomly selected male soccer players and concluded that frequently heading the ball in soccer lowers players' IQs. The players were divided into two groups based on how many headers they averaged per game. The following data was reported. Do these data support the researcher's conclusion at a 5% significance level? Can you conclude heading the ball causes lower IQ?

	n	Mean	Std Dev
< 10 Headers	35	112	10
≥ 10 Headers	25	103	8