#### MATH 450: Mathematical statistics

Oct 10th, 2019

Lecture 14: Large-sample CIs of the population mean

# Countdown to midterm: 14 days

Week 2 · · · · •	Chapter 6: Statistics and Sampling Distributions
Week 4 · · · ·	Chapter 7: Point Estimation
Week 7 · · · ·	Chapter 8: Confidence Intervals
Week 10 · · · ·	Chapter 9: Test of Hypothesis
Week 11 · · · · ·	Chapter 10: Two-sample inference
Week 13 · · · · ·	Regression

#### Overview

- 8.1 Basic properties of confidence intervals (CIs)
  - Interpreting Cls
  - General principles to derive CI
- 8.2 Large-sample confidence intervals for a population mean
  - Using the Central Limit Theorem to derive CIs
- 8.3 Intervals based on normal distribution
  - Using Student's t-distribution
- 8.4 Cls for standard deviation

#### Framework

- Let  $X_1, X_2, ..., X_n$  be a random sample from a distribution  $f(x, \theta)$
- $\bullet$  In Chapter 7, we learnt methods to construct an estimate  $\hat{\theta}$  of  $\theta$
- Goal: we want to indicate the degree of uncertainty associated with this random prediction
- One way to do so is to construct a *confidence interval*  $[\hat{\theta}-a,\hat{\theta}+b]$  such that

$$P[\theta \in [\hat{\theta} - a, \hat{\theta} + b]] = 95\%$$



## Principles for deriving CIs

If  $X_1, X_2, \ldots, X_n$  is a random sample from a distribution  $f(x, \theta)$ , then

- Find a random variable  $Y = h(X_1, X_2, ..., X_n; \theta)$  such that the probability distribution of Y does not depend on  $\theta$  or on any other unknown parameters.
- Find constants a, b such that

$$P[a < h(X_1, X_2, \dots, X_n; \theta) < b] = 0.95$$

ullet Manipulate these inequalities to isolate heta

$$P[\ell(X_1, X_2, \dots, X_n) < \theta < u(X_1, X_2, \dots, X_n)] = 0.95$$



## Confidence interval: example

#### Problem

Suppose the sediment density (g/cm) of a randomly selected specimen from a certain region is normally distributed with mean  $\mu$  and standard deviation 0.85.

If a random sample of 25 specimens is selected, with sample average  $\bar{X}$ .

• Find a number a such that

$$P[-a < \bar{X} - \mu < a] = 0.95$$

Confidence intervals for a population mean

#### 8.1: Normal distribution with know $\sigma$

- Assumptions:
  - Normal distribution
  - $\bullet$   $\sigma$  is known
- 95% confidence interval If after observing  $X_1 = x_1$ ,  $X_2 = x_2$ ,...,  $X_n = x_n$ , we compute the observed sample mean  $\bar{x}$ . Then

$$\left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

is a 95% confidence interval of  $\mu$ 



### **Assumptions**

- Section 8.1
  - Normal distribution
  - $\bullet$   $\sigma$  is known
- Section 8.2
  - Normal distribution
    - $\rightarrow$  use Central Limit Theorem  $\rightarrow$  needs n > 30
  - $\bullet$   $\sigma$  is known
    - $\rightarrow$  replace  $\sigma$  by  $s \rightarrow$  needs n > 40
- Section 8.3
  - Normal distribution
  - $\bullet$   $\sigma$  is known
  - $\rightarrow$  Introducing *t*-distribution

#### z-critical value

#### NOTATION

 $z_{\alpha}$  will denote the value on the measurement axis for which  $\alpha$  of the area under the z curve lies to the right of  $z_{\alpha}$ . (See Figure 4.19.)

For example,  $z_{.10}$  captures upper-tail area .10 and  $z_{.01}$  captures upper-tail area .01.

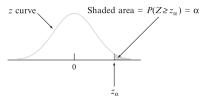


Figure 4.19  $z_{\alpha}$  notation illustrated

Since  $\alpha$  of the area under the standard normal curve lies to the right of  $z_{\alpha}$ ,  $1-\alpha$  of the area lies to the left of  $z_{\alpha}$ . Thus  $z_{\alpha}$  is the  $100(1-\alpha)$ th percentile of the standard normal distribution. By symmetry the area under the standard normal curve to the left of  $-z_{\alpha}$  is also  $\alpha$ . The  $z_{\alpha}$ 's are usually referred to as z critical values. Table 4.1 lists the most useful standard normal percentiles and  $z_{\alpha}$  values.

## $100(1-\alpha)\%$ confidence interval

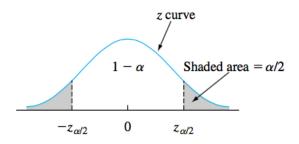


Figure 8.4  $P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1 - \alpha$ 

## $100(1-\alpha)\%$ confidence interval

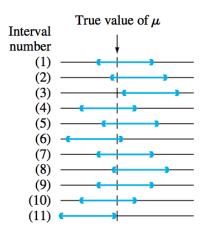
A  $100(1-\alpha)\%$  confidence interval for the mean  $\mu$  of a normal population when the value of  $\sigma$  is known is given by

$$\left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) \tag{8.5}$$

or, equivalently, by  $\bar{x} \pm z_{\alpha/2} \cdot \sigma / \sqrt{n}$ .

									X-7	
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997

## Interpreting confidence intervals



95% confidence interval: If we repeat the experiment many times, the interval contains  $\mu$  about 95% of the time

## Interpreting confidence intervals

Writing

$$P[\mu \in (\bar{X} - 1.7, \bar{X} + 1.7)] = 95\%$$

is okay.

• If  $\bar{x} = 2.7$ , writing

$$P[\mu \in (1, 4.4)] = 95\%$$

is NOT okay.

- Saying  $\mu \in (1, 4.4)$  with confidence level 95% is okay.
- Saying "if we repeat the experiment many times, the interval contains  $\mu$  about 95% of the time" is perfect.

## Example 1

#### Example

Assume that the helium porosity (in percentage) of coal samples taken from any particular seam is normally distributed with true standard deviation  $\sigma=.75$ .

- Compute a 95% CI for the true average porosity of a certain seam if the average porosity for 20 specimens from the seam was 4.85.
- How large a sample size is necessary if the width of the 95% interval is to be .40?

Review: sample variance

### Measures of Variability: deviations from the mean

Given a data set  $x_1, x_2, \ldots, x_n$ :

The sample variance, denoted by  $s^2$ , is given by

$$s^{2} = \frac{\sum (x_{i} - \bar{x})^{2}}{n - 1} = \frac{S_{xx}}{n - 1}$$

The **sample standard deviation**, denoted by s, is the (positive) square root of the variance:

$$s = \sqrt{s^2}$$

# Computing formula for $s^2$

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}$$

Proof Because 
$$\bar{x} = \sum x_i / n$$
,  $n\bar{x}^2 = (\sum x_i)^2 / n$ . Then,  

$$\sum (x_i - \bar{x})^2 = \sum (x_i^2 - 2\bar{x} \cdot x_i + \bar{x}^2) = \sum x_i^2 - 2\bar{x} \sum x_i + \sum (\bar{x})^2$$

$$= \sum x_i^2 - 2\bar{x} \cdot n\bar{x} + n(\bar{x})^2 = \sum x_i^2 - n(\bar{x})^2$$

8.2: Large-sample CIs of the population mean

## **Principles**

Central Limit Theorem

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

is approximately normal when n > 30

- Moreover, when *n* is sufficiently large  $s \approx \sigma$
- Conclusion:

$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$

is approximately normal when n is sufficiently large

If n>40, we can ignore the normal assumption and replace  $\sigma$  by s

### 95% confidence interval

If after observing  $X_1=x_1,\ X_2=x_2,\ldots,\ X_n=x_n\ (n>40)$ , we compute the observed sample mean  $\bar{x}$  and sample standard deviation s. Then

$$\left(\bar{x} - 1.96 \frac{s}{\sqrt{n}}, \bar{x} + 1.96 \frac{s}{\sqrt{n}}\right)$$

is a 95% confidence interval of  $\mu$ 

## $100(1-\alpha)\%$ confidence interval

If after observing  $X_1 = x_1$ ,  $X_2 = x_2$ ,...,  $X_n = x_n$  (n > 40), we compute the observed sample mean  $\bar{x}$  and sample standard deviation s. Then

$$\left(\bar{x}-z_{\alpha/2}\frac{s}{\sqrt{n}},\bar{x}+z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

is a 95% confidence interval of  $\mu$ 

One-sided CIs (Confidence bounds)

#### One-sided Cls

A large-sample upper confidence bound for  $\mu$  is

$$\mu < \bar{x} + z_{\alpha} \cdot \frac{s}{\sqrt{n}}$$

and a large-sample lower confidence bound for  $\mu$  is

$$\mu > \bar{x} - z_{\alpha} \cdot \frac{s}{\sqrt{n}}$$

#### Cls vs. one-sided Cls

Cls:

•  $100(1-\alpha)\%$  confidence

$$\left(\bar{x}-z_{\alpha/2}\frac{s}{\sqrt{n}},\bar{x}+z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

• 95% confidence

$$\left(\bar{x} - 1.96 \frac{s}{\sqrt{n}}, \bar{x} + 1.96 \frac{s}{\sqrt{n}}\right)$$

One-sided Cls:

•  $100(1-\alpha)\%$  confidence

$$\left(-\infty, \bar{x} + z_{\alpha} \frac{s}{\sqrt{n}}\right)$$

95% confidence

$$\left(-\infty, \bar{x} + 1.64 \frac{s}{\sqrt{n}}\right)$$

#### Confidence level

#### Problem

Determine the confidence level for each of the following large-sample confidence intervals/bounds:

(a) 
$$\bar{x} + 0.84s / \sqrt{n}$$

(b) 
$$(\bar{x} - 0.84s/\sqrt{n}, \bar{x} + 0.84s/\sqrt{n})$$

(c) 
$$\bar{x} - 2.05s/\sqrt{n}$$

									X-7	
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997

## Example 2

#### Example

A sample of 66 obese adults was put on a low-carbohydrate diet for a year. The average weight loss was 11 lb and the standard deviation was 19 lb. Calculate a 99% lower confidence bound for the true average weight loss

8.3: Intervals based on normal distributions

## Assumptions

- the population of interest is normal (i.e.,  $X_1, \ldots, X_n$  constitutes a random sample from a normal distribution  $\mathcal{N}(\mu, \sigma^2)$ ).
- $\sigma$  is unknown
- $\rightarrow$  we want to consider cases when *n* is small.

# **Principles**

• When n < 40, S is no longer close to  $\sigma$ . Thus

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

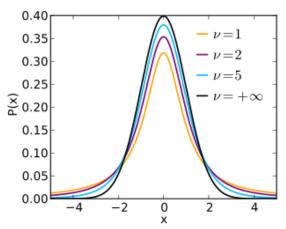
does not follow the standard normal distribution.

- {Section 6} But since we know the distribution of *X*, technically we can compute the distribution of *T*
- Moreover, the distribution of T does not depend on  $\mu$  and  $\sigma$  {More reading: Section 6.4}

### t distributions with degree of freedom $\nu$

#### Probability density function

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$



#### t distributions

#### PROPERTIES OF T DISTRI-BUTIONS

- 1. Each  $t_v$  curve is bell-shaped and centered at 0.
- **2.** Each  $t_v$  curve is more spread out than the standard normal (z) curve.
- 3. As v increases, the spread of the  $t_v$  curve decreases.
- **4.** As  $v \to \infty$ , the sequence of  $t_v$  curves approaches the standard normal curve (so the z curve is often called the t curve with df =  $\infty$ ).

#### t distributions

When  $\bar{X}$  is the mean of a random sample of size n from a normal distribution with mean  $\mu$ , the rv

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has the t distribution with n-1 degree of freedom (df).

#### t distributions

Let  $t_{\alpha,\nu}$  = the number on the measurement axis for which the area under the t curve with  $\nu$  df to the right of  $t_{\alpha,\nu}$ , is  $\alpha$ ;  $t_{\alpha,\nu}$  is called a t critical value.

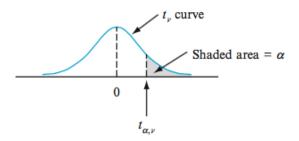


Figure 8.7 A pictorial definition of  $t_{\alpha,\nu}$ 

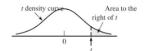
### How to do computation with t distributions

- Instead of looking up the normal Z-table A3, look up the two t-tables A5 and A7.
- Idea

$$P[T \ge t_{\alpha,\nu}] = \alpha$$

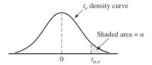
- {From t, find  $\alpha$ }  $\rightarrow$  using table A7
- {From  $\alpha$ , find t}  $\rightarrow$  using table A5

**Table A.7** t Curve Tail Areas



t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0.0	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
0.1	.468	.465	.463	.463	.462.	.462	.462	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461
0.2	.437	.430	.427	.426	.425	.424	.424	.423	.423	.423	.423	.422	.422	.422	.422	.422	.422	.422
0.3	.407	.396	.392	.390	.388	.387	.386	.386	.386	.385	.385	.385	.384	.384	.384	.384	.384	.384
0.4	.379	.364	.358	.355	.353	.352	.351	.350	.349	.349	.348	.348	.348	.347	.347	.347	.347	.347
0.5	.352	.333	.326	.322	.319	.317	.316	.315	.315	.314	.313	.313	.313	.312	.312	.312	.312	.312
0.6	.328	.305	.295	.290	.287	.285	.284	.283	.282	.281	.280	.280	.279	.279	.279	.278	.278	.278
0.7	.306	.278	.267	.261	.258	.255	.253	.252	.251	.250	.249	.249	.248	.247	.247	.247	.247	.246
0.8	.285	.254	.241	.234	.230	.227	.225	.223	.222	.221	.220	.220	.219	.218	.218	.218	.217	.217
0.9	.267	.232	.217	.210	.205	.201	.199	.197	.196	.195	.194	.193	.192	.191	.191	.191	.190	.190
1.0	.250	.211	.196	.187	.182	.178	.175	.173	.172	.170	.169	.169	.168	.167	.167	.166	.166	.165
1.1	.235	.193	.176	.167	.162	.157	.154	.152	.150	.149	.147	.146	.146	.144	.144	.144	.143	.143
1.2	.221	.177	.158	.148	.142	.138	.135	.132	.130	.129	.128	.127	.126	.124	.124	.124	.123	.123
1.3	.209	.162	.142	.132	.125	.121	.117	.115	.113	.111	.110	.109	.108	.107	.107	.106	.105	.105
1.4	.197	.148	.128	.117	.110	.106	.102	.100	.098	.096	.095	.093	.092	.091	.091	.090	.090	.089
1.5	.187	.136	.115	.104	.097	.092	.089	.086	.084	.082	.081	.080	.079	.077	.077	.077	.076	.075
1.6	.178	.125	.104	.092	.085	.080	.077	.074	.072	.070	.069	.068	.067	.065	.065	.065	.064	.064
1.7	.169	.116	.094	.082	.075	.070	.065	.064	.062	.060	.059	.057	.056	.055	.055	.054	.054	.053
1.8	.161	.107	.085	.073	.066	.061	.057	.055	.053	.051	.050	.049	.048	.046	.046	.045	.045	.044
1.9	.154	.099	.077	.065	.058	.053	.050	.047	.045	.043	.042	.041	.040	.038	.038	.038	.037	.037

**Table A.5** Critical Values for *t* Distributions



	α											
ν	.10	.05	.025	.01	.005	.001	.0005					
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62					
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598					
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924					
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610					
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869					
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959					
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408					
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041					
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781					
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587					
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437					
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318					
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221					
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140					
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073					
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015					
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965					

#### Confidence intervals

Let  $\bar{x}$  and s be the sample mean and sample standard deviation computed from the results of a random sample from a normal population with mean  $\mu$ . Then a  $100(1 - \alpha)\%$  confidence interval for  $\mu$ , the one-sample t CI, is

$$\left(\overline{x} - t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}}, \overline{x} + t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}}\right) \tag{8.15}$$

or, more compactly,  $\bar{x} \pm t_{\alpha/2,n-1} \cdot s/\sqrt{n}$ .

An upper confidence bound for  $\mu$  is

$$\bar{x} + t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}}$$

and replacing + by – in this latter expression gives a **lower confidence** bound for  $\mu$ ; both have confidence level  $100(1 - \alpha)\%$ .

# Example 3

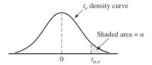
#### Example

Here is a sample of ACT scores for students taking college freshman calculus:

24.00	28.00	27.75	27.00	24.25	23.50	26.25
24.00	25.00	30.00	23.25	26.25	21.50	26.00
28.00	24.50	22.50	28.25	21.25	19.75	

Assume that ACT scores are normally distributed, calculate a two-sided 95% confidence inter-val for the population mean.

**Table A.5** Critical Values for *t* Distributions



	α											
ν	.10	.05	.025	.01	.005	.001	.0005					
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62					
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598					
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924					
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610					
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869					
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959					
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408					
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041					
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781					
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587					
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437					
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318					
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221					
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140					
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073					
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015					
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965					