

20. CONFIDENCE INTERVALS FOR THE MEAN, UNKNOWN VARIANCE

If the population standard deviation σ is unknown, as it usually will be in practice, we will have to estimate it by the sample standard deviation s .

Since σ is unknown, we cannot use the confidence intervals described previously. The practical versions presented here use s in place of σ .

Eg 1: A random sample of 8 “Quarter Pounders” yields a mean weight of $\bar{x} = 0.2$ pounds, with a sample standard deviation of $s = 0.07$ pounds. Construct a 95% CI for the unknown population mean weight for all “Quarter Pounders”. (Solution on next slide).

Remember that s is the square root of the sample variance,

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{where } x_1 \cdots x_n \text{ are the data values in the sample.}$$

There are two cases to consider: “Large Sample” and “Small Sample”.

- If $n \geq 30$ (and σ is unknown), just replace σ by s and proceed as before. The CI is $\bar{X} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$.

This works because $\frac{\bar{X} - \mu}{s / \sqrt{n}}$ is approximately standard normal,

regardless of the population distribution, when $n \geq 30$.

We can think of $\frac{s}{\sqrt{n}}$ as the *estimated* standard error of \bar{X} .

- Suppose $n < 30$ (and σ is unknown). To get a valid CI for μ in this case, we must assume that the population distribution is normal. This assumption is hard to check, and was not required before.

The CI is $\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$, where $t_{\alpha/2}$ is defined below.

- Don't forget what you learned earlier: If σ is known, we use $\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$, regardless of the sample size. So the only time when you must remember to use $t_{\alpha/2}$ is when σ is unknown and $n < 30$. Otherwise, use $z_{\alpha/2}$.

SUMMARY OF CONFIDENCE INTERVALS

	Sigma Known	Sigma Unknown
$n \geq 30$	$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	$\bar{X} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$
$n < 30$	$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	$\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

* Must assume Normal population if $n < 30$, σ unknown.

What is $t_{\alpha/2}$, and why must we use it when $n < 30$ and σ is unknown?

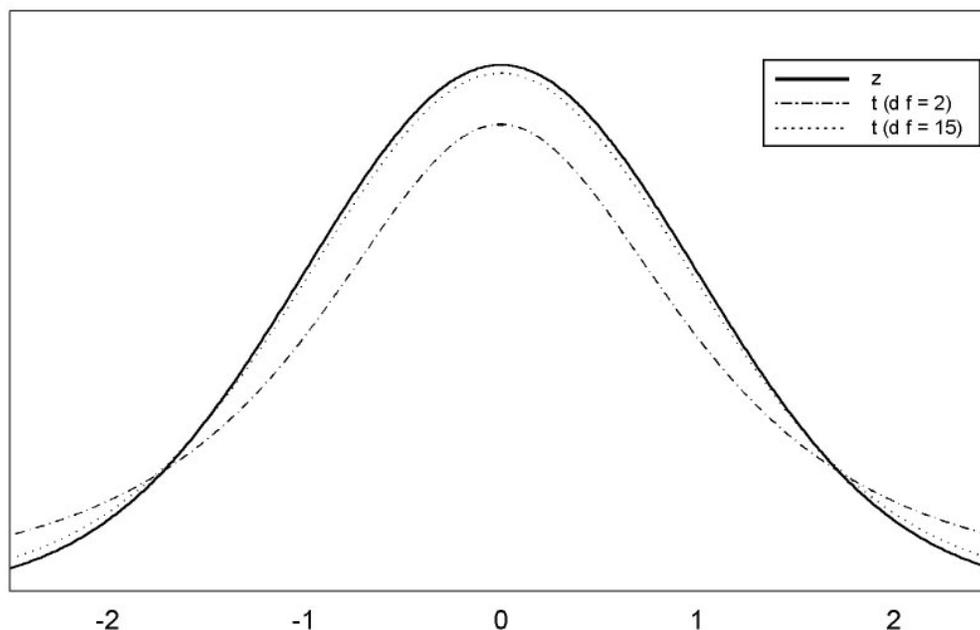
When $n < 30$, the quantity $t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$ does not have an approximately standard normal distribution, even though we assume here that the population is normal.

Instead, t has a “Student’s t distribution with $n - 1$ degrees of freedom”.

This distribution was invented by W.S. Gosset (1908), who was working for Guinness Breweries, and wrote under the pen name of “Student”.

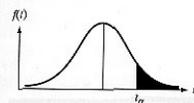
There is a different t distribution for each value of the degrees of freedom, df . These distributions are not normal, although they are symmetrical around zero, and mound shaped.

Standard Normal and t Distributions



The quantity $t_{\alpha/2}$ denotes the t -value such that the area to its right under the Student's t distribution (with $df = n - 1$) is $\alpha/2$. Note that we use $n - 1$ df , even though the sample size is n . Values of t_{α} are listed in Table VI, Appendix B. In the table, the degrees of freedom are denoted by v .

Table 6 Critical Values for Student's t



v	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
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
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: This table is reproduced with the kind permission of the Trustees of Biometrika from E. S. Pearson and H. O. Hartley (eds.), *The Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., Biometrika, 1966.

Eg 1, Solution: We have $n = 8$, $\bar{x} = 0.2$, $s = 0.07$, $\alpha = 0.05$, $df = n - 1 = 7$. From Table VI, $t_{\alpha/2} = t_{0.025} = 2.365$. Therefore, the confidence interval is

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 0.2 \pm 2.365 \frac{0.07}{\sqrt{8}} = 0.2 \pm 0.059 = (0.141, 0.259).$$

It's interesting that this interval contains 0.25, so it seems plausible that the burgers are really quarter pounders (on average).

Relationship Between t and z

When df is small (the guideline we're using is $df < 29$), the t distribution has “longer tails” (i.e., contains more outliers) than the normal distribution, and it is important to use the t -values of Table VI, assuming that σ is unknown.

Due to the long tails in the t distribution, $t_{\alpha/2}$ is larger than $z_{\alpha/2}$, so the confidence interval based on t will be wider than the (incorrect) one based on z .

So when df is small, the t -based confidence interval correctly reflects the added uncertainty due to not knowing σ , which must be estimated.

As df gets large, however, the t distribution approaches the standard normal distribution.

Therefore, if n is large enough, $t_{\alpha/2}$ will be extremely close to $z_{\alpha/2}$, and the two confidence intervals

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \quad \text{and} \quad \bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

will be virtually identical, so it will not matter which one we use.

But how large is “large”?

The Traditional Approach (in keeping with the text):

Any value of $df \geq 29$ is considered “infinite”. We can then use the ∞ row of Table VI, which gives z_{α} (that is, t_{α} with $df = \infty$).

Minitab’s Approach: The df is never truly infinite, so Minitab always uses $t_{\alpha/2}$ in computing confidence intervals. (Minitab has its own “internal tables” to calculate $t_{\alpha/2}$ for any df).

To be consistent with the textbook, and because our table does not have listings for all values of df, we will use the traditional approach whenever we calculate confidence intervals by hand.

Minitab's approach is actually the correct one, however, and is to be preferred whenever $t_{\alpha/2}$ and $z_{\alpha/2}$ differ appreciably.

Eg 2: Minitab Descriptive Statistics for Chips Ahoy! data:

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Chips	42	1261.6	1241.5	1255.5	117.6	18.1

Let's construct a 99% confidence interval for μ , the mean number of chips per bag. The Traditional Approach gives

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = 1261.6 \pm 2.576 (117.6) / \sqrt{42} = (1214.9, 1308.3) .$$

Minitab's 1-Sample t output gives a wider interval, since it uses the exact value, $t_{.005} = 2.701$ for $df = 41$:

T Confidence Intervals

Variable	N	Mean	StDev	SE Mean	99.0 % CI
Chips	42	1261.6	117.6	18.1	(1212.6, 1310.6)

In the Sampling Lab, each of you constructed your own small-sample confidence interval for the mean. Since we know the mean for our population, we can see what percentage of the intervals actually “worked”.

[Sampling Lab Results].

Eg 3: The United States Chess Federation (USCF) has a rating system, which they use to predict the expected outcome of tournament games. Each player has a rating between 0 and 3,000 which changes over time. The larger the difference in ratings between two players, the greater the probability that the higher rated player will win. A chess game is scored 0 for a loss, 1/2 for a draw, and 1 for a win. Figure 3 uses confidence intervals to show that the predicted winning expectancy based on the difference in USCF ratings is quite far from the actual average outcome. In other words, the USCF ratings don’t work!

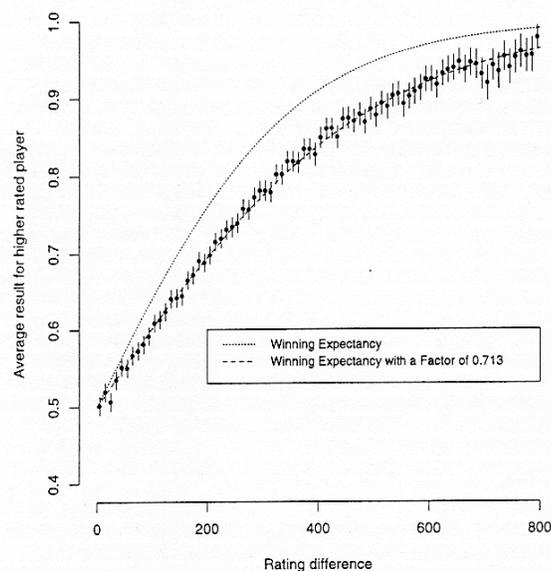


Figure 3. Summary of 225,621 rated USCF tournament games. Both players in a game must have had established ratings in 1997 to be included in the sample. The sample is partitioned into groups of players according to their rating difference (0–10, 10–20, ..., 790–800). For each rating difference group, the dot represents the average score of games relative to the higher-rated player. The vertical bars represent 95% confidence intervals. The values on the dotted line are the expected scores calculated from the USCF winning expectancy formula. The values on the dashed line (which trace through the segments) represent a better-fitting model.