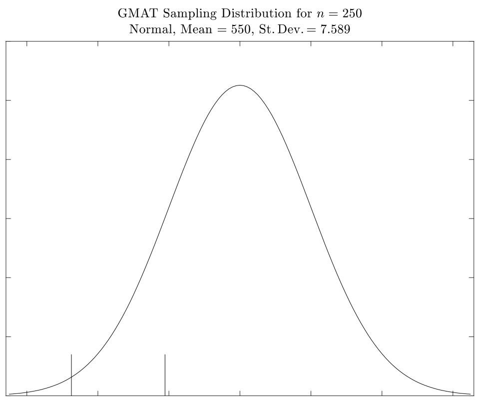
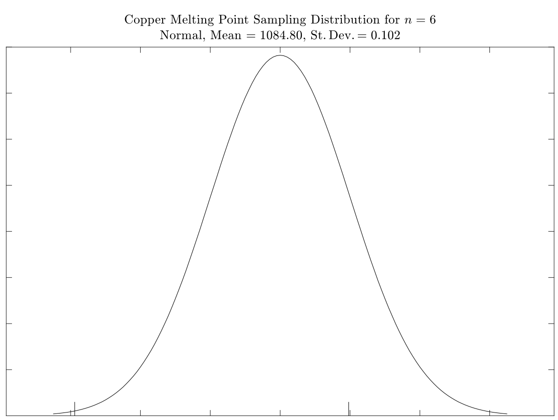
Chapter 17 – Tests of Significance: The Basics

17.1 (a) If μ = 550, the sampling distribution is approximately Normal, with mean μ = 550 and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{120}{\sqrt{250}} = 7.59$. The density curve is provided.



(b) and **(c)** Both points are marked on the figure in part (a). A sample mean $\bar{x} = 542$ lies just slightly further than 1 standard deviation below the mean, while $\bar{x} = 532$ lies toward the low tail of the curve. If μ = 550, observing a value of 542 is not too surprising, but observing a value of 532 is much less likely, which provides some evidence that μ < 550.

17.2 (a) If μ = 1084.80, the sampling distribution is approximately Normal, with mean μ = 1084.80 and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{0.25}{\sqrt{6}} = 0.102$. The density curve is provided.



(b) Both points are marked on the figure in part (a). A sample mean $\bar{x} = 1084.90$ lies fairly close to the middle of the distribution, while $\bar{x} = 1084.50$ lies nearly 3 standard deviations below the mean. If μ = 1084.80, observing a value of 1084.90 is not too surprising, but observing a value of 1084.50 is much less likely, which provides some evidence that $\mu \neq 1084.80$°C.

17.3 $H_0: \mu = 550$ vs. $H_a: \mu < 550$. Part (c) of Exercise 17.1 refers to providing evidence that the mean score is less than 550, so we need to use the one–sided test with the less than alternative hypothesis.

17.4 $H_0: \mu = 1084.80$ vs. $H_a: \mu \neq 1084.80$. Part (b) of Exercise 17.2 refers to providing evidence that the true melting point differs from 1084.80°C, so we need to use the two–sided test.

17.5 $H_0: \mu = 75$ vs. $H_a: \mu < 75$. The professor suspects that this section's students perform worse than the population of all students in the class, on average.

17.6 $H_0: \mu = \$35{,}713$ vs. $H_a: \mu \neq \$35{,}713$. This is a two–sided test, because you wonder if the full–time income for women high school graduates in your school differs from the national average.

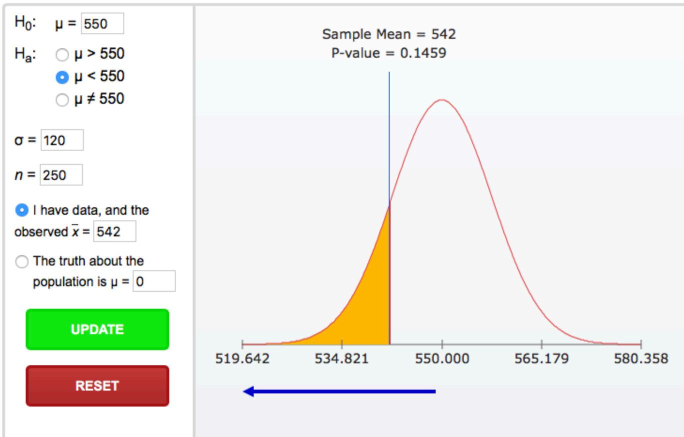
17.7 Hypotheses are statements about parameters, not statistics. The research question should not be about the sample mean ($\bar{x}$), but should be about the population mean, μ.

17.8 (a) With σ = 1 and n = 10, the standard deviation is $\frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{10}} = 0.3162$, so when μ = 0, the distribution of $\bar{x}$ is $N(0, 0.3162)$. **(b)** The P–value is $P = P(\bar{x} \geq 0.3) = P(Z \geq \frac{0.3-0}{0.3162}) = 0.1711$.

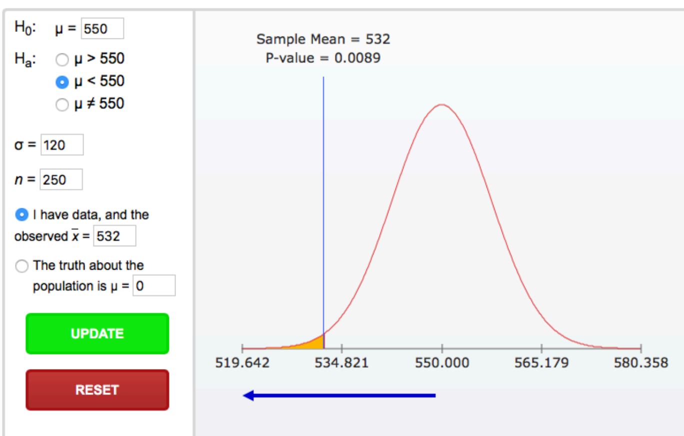
17.9 (a) With σ = 60 and n = 18, the standard deviation is $\frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{18}} = 14.1421$, so when μ = 0, the distribution of $\bar{x}$ is $N(0, 14.1421)$. **(b)** The P–value is $P = 2P(\bar{x} \geq 17) = 2P(Z \geq \left| \frac{17-0}{14.1421} \right|) = 0.2302$.

17.10 If the drug lorcaserin had no effect, then we would expect the average weight loss for the treatment group (on lorcaserin) to be the same as the average weight loss for the control group (on placebo). Of course, by random chance alone, there will be some difference in these average weight losses. The P–value $P < 0.001$ says that if lorcaserin is ineffective, then the chance of observing as large an increased weight loss as the difference observed (5.8 kg − 2.2 kg = 3.6 kg) would happen by chance alone less than 1 in 1000 times. That is, we have seen something that we do not expect to occur by random chance, and this causes us to doubt the original assumption that lorcaserin is ineffective. Random chance alone does not explain the additional weight loss in the lorcaserin group.

17.11 (a) One realization from the applet is provided. The P–value for $\bar{x}$ = 542 is 0.1459. (Calculating the P–value by hand with rounding similar to other exercises, one finds P–value = 0.1469.) The P–value is not significant at either α = 0.05 or α = 0.01.

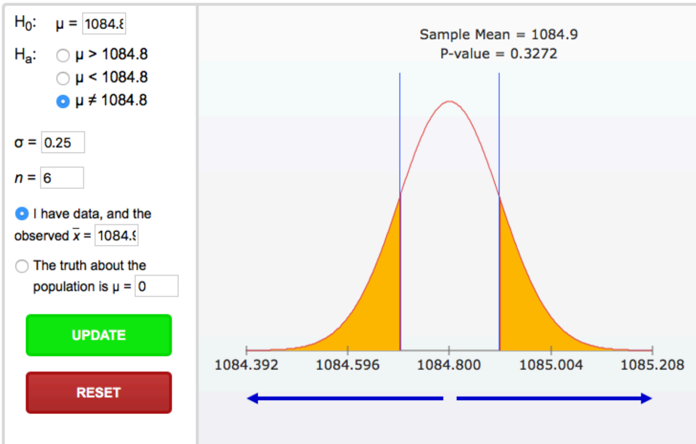


(b) One realization from the applet is provided. The P–value for $\bar{x}$ = 532 is 0.0089. This is significant at both α = 0.05 and α = 0.01.

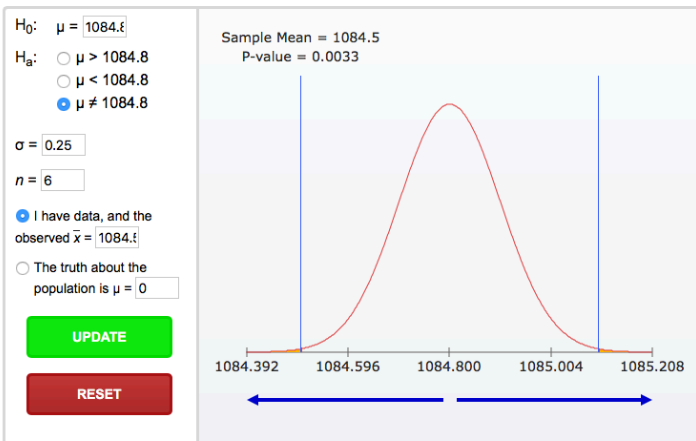


(c) If μ = 550 (that is, if H_0 were true), observing a value similar to 542 would not be too surprising, but observing 532 is not very likely at all, which provides strong evidence that μ < 550.

17.12 (a) One realization from the applet is provided. The P–value for $\bar{x}$ = 1084.90 is 0.3272. This is not significant at either α = 0.05 or α = 0.01.



(b) One realization from the applet is provided. The *P*–value for $\bar{x}$ = 1084.50 is 0.0033. This is significant at both α = 0.05 and α = 0.01.



(c) If μ = 1084.80 (that is, if H_0 were true), observing a value similar to 1084.90 would not be too surprising. Observing 1084.50 is much less likely; this extreme observation provides strong evidence that $\mu \neq 1084.80$°C.

17.13 **(a)** $z = \frac{0.3 - 0}{1/\sqrt{10}} = \frac{0.3 - 0}{0.3162} = 0.9488$. **(b)** $z = \frac{1.02 - 0}{1/\sqrt{10}} = \frac{1.02 - 0}{0.3162} = 3.226$. **(c)** $z = \frac{17 - 0}{60/\sqrt{18}} = \frac{17 - 0}{14.1421} = 1.2021$. Note that in part (c) the test is two–sided, while in parts (a) and (b), it is one–sided.

17.14 STATE: Is there evidence that the true melting point of the copper sample is not 1084.80°C? PLAN: Let μ be the sample's true melting point (the mean of all measurements of its melting point). We test $H_0: \mu = 1084.80$ vs. $H_a: \mu \neq 1084.80$ using the two–sided alternative, because we are concerned with deviations in either direction. SOLVE: The problem states that we have an SRS from a Normal distribution. From the data, $\bar{x} = 1084.80$. The standard deviation of $\bar{x}$ is $\frac{\sigma}{\sqrt{n}} = \frac{0.25}{\sqrt{6}} = 0.1021$, so the test statistic is $z = \frac{1084.80 - 1084.80}{0.1021} = 0$. The *P*–value is $2P(Z \geq 0) = 1$.

CONCLUDE: This sample gives absolutely no evidence that the true melting point of the copper sample differs from 1084.80°C.

17.15 STATE: Is there evidence that the average tip percentage is less than 20% when bad news is received (such as a bad weather prediction)? PLAN: Let μ be the average tip percentage for all customers receiving bad news. We test H_0: μ = 20 against H_a: μ < 20, since we wonder if the value of μ is less than 20%. SOLVE: We have a random sample of n = 20 customers and were told to assume tips have a Normal distribution. We observe $\bar{x}$ = 18.19%. The standard deviation of $\bar{x}$ is $\frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{20}} = 0.4472$, so the test statistic is $z = \frac{18.19 - 20}{0.4472} = -4.05$. The P–value is $P(Z \leq -4.05) \approx 0$. CONCLUDE: There is overwhelming evidence that the average tip percentage, when bad news is delivered, is lower than the average tip percentage overall. Random chance does not explain the small value of $\bar{x}$ observed.

17.16 Using Table C, z = 1.65 is significant at α = 0.05, because it is larger than 1.645. It is not significant at α = 0.01, because it is smaller than 2.326.

17.17 Using Table C, z = 1.65 is not significant at α = 0.05, because it is not larger than 1.960 or smaller than −1.960. It is also not significant at α = 0.01, because $|z|$ is smaller than 2.576. (If z is not significant at a particular α, it will also be not significant at a smaller α.)

17.18 (a) $z = \frac{0.4365 - 0.5}{0.2887/\sqrt{100}} = -2.20$. **(b)** This result is significant at the 5% level, because z < −1.96. **(c)** It is not significant at the 1% level, because $-2.576 < z < 2.576$. **(d)** This value of z is between −2.054 and −2.326, so the P–value is between 0.02 and 0.04 (because the alternative is two–sided).

17.19 (a) true, of the test statistic taking a value as extreme as or more extreme than that actually observed is 0.011. This is the definition of a P–value.

17.20 (b) statistically significant at $\alpha = 0.05$ but not at $\alpha = 0.01$. P = 0.011 is less than 0.05, but not 0.01.

17.21 (b) statistically significant at $\alpha = 0.05$ but not at $\alpha = 0.01$. The P–value for z = 2.29 is 0.0110 (assuming that the difference is in the correct direction; that is, assuming that the alternative hypothesis was H_a: $\mu > \mu_0$).

17.22 (b) z = −0.577. $z = \frac{19.667 - 20}{1/\sqrt{3}} = -0.577$.

17.23 (a) H_0: $\mu = 18$. The null hypothesis states that μ takes on the "default" value, 18 seconds.

17.24 (b) H_a: $\mu < 18$. The researcher believes that loud noises will make the rats complete the maze faster (decrease the completion time), so the alternative is one–sided.

17.25 (c) neither of the above is true. A small *P*–value means we should not (or should rarely) find an observed difference as large or larger than what was seen in H_0 is true. The *P*–value does not tell us whether the difference seen is "large" or "practically important," nor does it refer to the probability H_0 is true.

17.26 (c) All values for which $|z| > 2.807$. This is a two–sided alternative, so we have 0.0025 in each tail of the Normal distribution, leading to $|z| > 2.807$.

17.27 (a) All values for which $z > 2.576$. This is a one–sided alternative, so we have 0.005 in the right tail of the Normal distribution, leading to $z > 2.576$.

17.28 (a) H_0: $\mu = 13$ hours per week vs. H_a: $\mu > 13$ hours per week. **(b)** $z = \frac{13.7 - 13}{7.4/\sqrt{463}} = 2.04$. **(c)** *P*–value = $P(Z > 2.04) = 0.0207$. There is strong evidence that students do claim to study more than 13 hours per week on the average.

17.29 (a) We test H_0: $\mu = 0$ vs. H_a: $\mu > 0$. **(b)** $z = \frac{2.35 - 0}{2.5/\sqrt{200}} = 13.29$. **(c)** This value of z is far outside the range we would expect from the $N(0, 1)$ distribution. Under H_0, it would be virtually impossible to observe a sample mean as large as 2.35 based on a sample of 200 men. The sample mean is not explained by random chance, and we would easily reject H_0.

17.30 (a) We test H_0: $\mu = 5.19$ against H_0: $\mu \neq 5.19$. The alternative is two–sided, because we had no prior belief about the direction of the difference. (That is, before looking at the data, we had no reason to expect that the mean for hotel managers would be either higher or lower than 5.19.) **(b)** With $\bar{x} = 5.29$, the test statistic is $z = \frac{5.29 - 5.19}{0.78/\sqrt{148}} = 1.56$. **(c)** The *P*–value is $2P(Z > 1.56) = 0.1188$. There is only weak evidence that hotel managers have a different mean femininity score than the general male population. Particularly when the large sample ($n = 148$) is taken into account, we suspect that male hotel managers don't differ much from males in general (in this respect).

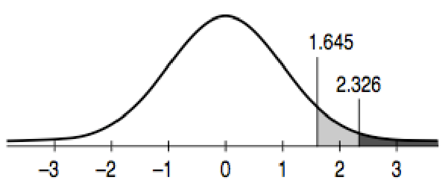
17.31 "$P = 0.005$" means that H_0 is not likely to be correct, but only in the sense that it provides a poor explanation of the data observed. It means that if H_0 is true, a sample as contrary to H_0 as our sample would occur by chance alone about 0.5% of the time, if the experiment was repeated over and over. However, it does not mean that there is a 0.5% chance that H_0 is true.

17.32 If the presence of pig skulls were not an indication of wealth, then differences similar to or bigger than those observed in this study would occur less than 1% of the time by chance.

17.33 The person making the objection is confusing practical significance with statistical significance. In fact, a 5% increase isn't a lot in a pragmatic sense. However, P = 0.03 means that random chance does not easily explain the difference observed. That is, there does seem to be an increase in mean improvement for those who expressed their anxieties, but the significance test does not address whether the difference is large enough to matter. Statistical significance is not practical significance.

17.34 With P = 0.24, any difference in quiz scores "caused" by instructor sex is easily explained by random chance. In other words, observing a difference of this magnitude was about the same as observing two heads when tossing two fair coins.

17.35 In the provided sketch, the "significant at 1%" region includes only the dark shading (z > 2.326). The "significant at 5%" region of the sketch includes both the light and dark shading (z > 1.645). When a test is significant at the 1% level, it means that if the null hypothesis were true, outcomes similar to (or more extreme than) those seen are expected in 0 or 1 of 100 repetitions of the experiment. When a test is significant at the 5% level, it means that if the null hypothesis were true, outcomes similar to (or more extreme than) those seen are expected in 5 or fewer of 100 repetitions of the experiment. Significance at the 1% level implies significance at the 5% level (or at any level higher than 1%). The converse is false; something that occurs "5 or fewer times in 100 repetitions" is not necessarily as rare as something that happens "1 or fewer times in 100 repetitions," so a test that is significant at the 5% level is not necessarily significant at the 1% level. Any z test statistic between 1.645 and 2.326 will be significant at the 5% level, but not at the 1% level.



17.36 (a) The researchers selected the alternative hypothesis after examining the data. The alternative hypothesis should be formulated before examining data, and especially should not be motivated by data. **(b)** The correct P–value is $2P(Z > 1.88)$ = 2(0.0301) = 0.0602.

17.37 Because a P–value is a probability, it can never be greater than 1. The correct P–value is $P(Z \geq 1.33) = 0.0918$.

17.38 (a) STATE: Can we conclude that the mean strength μ of wood pieces differs from 32,500 pounds? PLAN: We test H_0: μ = 32,500 against H_a: $\mu \neq$ 32,500 at the α = 0.10 level of significance. SOLVE: The sample mean is $\bar{x}$ = 30,841 pounds. The test statistic is $z = \frac{30{,}841 - 32{,}500}{3{,}000/\sqrt{20}} = -2.47$. The P–value is $P = 2P(Z \leq -2.47) = 0.0136$. CONCLUDE: There is enough evidence (by far) at the $\alpha = 0.10$ level of significance to conclude that the wood's mean strength differs from 32,500 pounds. **(b)** STATE: Can we conclude that the mean strength μ of wood pieces differs from 31,500 pounds? PLAN: We test H_0: $\mu = 31{,}500$ against H_a: $\mu \neq 31{,}500$ at the $\alpha = 0.10$ level of significance. SOLVE: The test statistic is $z = \frac{30{,}841 - 31{,}500}{3{,}000/\sqrt{20}} = -0.98$. The P–value is $P = 2P(Z \leq -0.98) = 0.327$. CONCLUDE: There is not enough evidence at the $\alpha = 0.10$ level of significance to conclude that the wood's mean strength differs from 31,500 pounds. Random chance easily explains the sample mean's distance from 31,500 pounds, but not from 32,500 pounds.

17.39 STATE: What is the mean percent change μ in spinal mineral content of nursing mothers? PLAN: We will test the hypotheses H_0: $\mu = 0\%$ against H_a: $\mu < 0\%$. SOLVE: The sample mean is $\bar{x} = -3.587\%$. The test statistic is $z = \frac{-3.587 - 0}{2.5/\sqrt{47}} = -9.84$, and the P–value is $P(Z \leq -9.84) \approx 0$. CONCLUDE: There is overwhelming evidence that, on average, nursing mothers lose bone mineral.

17.40 STATE: Is there evidence that the mean DMS threshold for untrained tasters is greater than 25 μg/L? PLAN: We test H_0: $\mu = 25$ μg/L vs. H_a: $\mu > 25$ μg/L. SOLVE: We find that $\bar{x} = 29.4$ μg/L, and the test statistic is $z = \frac{29.4 - 25}{7/\sqrt{10}} = 1.99$, so the P –value is $P(Z > 1.99) = 0.0233$. CONCLUDE: This is strong evidence against H_0 (at the α = 0.05 level); we conclude that the untrained student's mean threshold is greater than 25 μg/L.

17.41 (a) We test H_0: $\mu = 0$ vs. H_a: $\mu > 0$, where μ is the mean sensitivity difference in the population. **(b)** STATE: Does eye grease have a significant impact on eye sensitivity? PLAN: We test the hypotheses stated in part (a). SOLVE: The mean of the 16 differences is $\bar{x} = 0.10125$, so the test statistic is $z = \frac{0.10125 - 0}{0.22/\sqrt{16}} = 1.84$. The one–sided P–value for this value of z is $P = 0.0329$. CONCLUDE: The sample gives significant evidence (at $\alpha = 0.05$) that eye grease increases sensitivity.

17.42 (a) The margin of error for 90% confidence is $1.645(15/\sqrt{72}) = 2.908$, so the interval is $128.07 \pm 2.908 = 125.16$ to 130.98. **(b)** The test statistic is $z = \frac{128.07 - 130}{15/\sqrt{72}} = -1.09$, for which the two–sided P–value is $P = 0.2757$, which is greater than 0.10. **(c)** The test statistic is $z = \frac{128.07 - 131}{15/\sqrt{72}} = -1.66$, for which the two–sided P–value is $P = 0.0969$, which is (barely) less than 0.10.

17.43 (a) Yes, because 6.4 does not fall in the 95% confidence interval, which is (6.5, 7.9). **(b)** No, because 6.6 falls in the 95% confidence interval.

17.44 and **17.45** are Web–based exercises.