9.24 How to sell a burger A fast-food chain wants to compare two ways of promoting a new burger (a turkey burger). One way uses a coupon available in the store. The other way uses a poster display outside the store. Before the promotion, their marketing research group matches 50 pairs of stores. Each pair has two stores with similar sales volume and customer demographics. The store in a pair that uses coupons is randomly chosen, and after a month-long promotion, the increases in sales of the turkey burger are compared for the two stores. The increase was higher for 28 stores using coupons and higher for 22 stores using the poster. Is this strong evidence to support the coupon approach, or could this outcome be explained by chance? Answer by performing all five steps of a two-sided significance test about the population proportion of times the sales would be higher with the coupon promotion.

n=50, x=28(using coupons), $\hat{p}$=28/50=0.56

1) **categorical, random, $np_0 \geq 15$, $n(1-p_0) \geq 15$.**
2) **H_0: p=0.50; H_a: p$\neq 0.50$**
3) **Z=(0.56-0.50)/ $\sqrt{(0.5)(1-0.5)/50} = 0.85$**
4) **p value=2*P(Z>0.85)=0.3954**
5) **p is not small, do not reject H_0.**

If H_0 were true, there is 0.3954 probability of getting a result at least as extreme as the value observed.

Do not reject H_0. It is plausible that the null hypothesis is correct, that the coupons do not lead to higher sales than do posters.

(This decision is possibly a type II error).

9.36 Too little or too much wine? Wine-pouring vending machines, previously available in Europe and international airports, have become popular in the last few years in the United States. They are even approved to dispense wine in some Walmart stores. The available pouring options are a 5-ounce glass, a 2.5-ounce half-glass, and a 1-ounce taste. When the machine is in statistical control (see Exercise 7.42), the amount dispensed for a full glass is 5.1 ounces. Four observations are taken each day, to plot a daily mean over time on a control chart to check for irregularities. The most recent day's observations were 5.05, 5.15, 4.95, and 5.11. Could the difference between the sample mean and the target value be due to random variation, or can you conclude that the true mean is now different from 5.1? Answer by showing the five steps of a significance test, making a decision using a 0.05 significance level.

n=4, $\bar{x}$=5.065, s=0.087.

1) **Random, normal population.**
2) **H_0: μ=5.1; H_a: $\mu \neq 5.1$**
3) **t=(5.065-5.1)/ $\left(\frac{0.087}{\sqrt{4}}\right) = -0.80$**
4) **df=3, p value > 2*P(t>1.638)** **note: 1.638 is a value found on *t*-table**
   > 2* 0.1
   Exact P-value reported by the software = 2*P(t<-0.80)=0.48
5) **p is not small, do not reject H_0.**

If H_0 were true, there is 0.48 probability of getting a result at least as extreme as the value observed. There is not enough evidence to support that the true mean differs from 5.1 ounces.

Do not reject H_0.

(This decision is possibly a type II error).

9.37 Selling a burger In Exercise 9.24, a fast-food chain compared two ways of promoting a turkey burger. In a separate experiment with 10 pairs of stores, the difference in the month's increased sales between the store that used coupons and the store with the outside poster had a mean of \$3000. Does this indicate a true difference between mean sales for the two advertising approaches? Answer by using the output shown to test that the population mean difference is 0, carrying out the five steps of a significance test. Make a decision using a 0.05 significance level.

```
Test of mu = 0 vs.mu not = 0
Variable    N        Mean        StDev     SE Mean
Sales      10        3000        4000      1264.91

Variable           95.0% CI        T          P

Sales          (138.8, 5861.2)   2.372     0.04177
```

n=10, $\bar{x}$=3000, s=4000.

1) **Random, normal population.**
2) **H_0: μ=0; H_a: $\mu \neq 0$**
3) **t=(3000-0)/ $\left(\frac{4000}{\sqrt{10}}\right) = 2.372$**
4) **df=9, p value < 2*P(t>2.262)** **note: 1.638 is a value found on *t*-table**

   < 2* 0.025

   Exact P-value reported by the software = 2*P(t>2.372)=0.042

5) **P < 0.05, reject H_0 at 0.05 significance level.**

(This decision is possibly a type I error).

If H_0 were true, there is 0.042 probability of getting a result at least as extreme as the value observed. There is sufficient evidence that coupons led to higher sales than did the outside posters.

9.54 Effect of *n* Example 11 analyzed political conservatism and liberalism in the United States. Suppose that the sample mean of 4.11 and sample standard deviation of 1.43 were from a sample size of only 25, rather than 1933.

a. Find the test statistic.

b. Find the P-value for testing H_0: $\mu = 4.0$ against H_a: $\mu \neq 4.0$. Interpret.

c. Show that a 95% confidence interval for μ is (3.5, 4.7).

d. Together with the results of n=1933, explain what this illustrates about the effect of sample size on (i) the size of the P-value (for a given mean and standard deviation) and (ii) the width of the confidence interval.

$\bar{x} = 4.11$, **s=1.43**	n=25	n=1933
t =($\bar{x} - \mu_0$)/($\frac{s}{\sqrt{n}}$)	0.3846	3.38199
p-value	2*P(t(24)>0.3846) >2*P(t(24)>1.318) > 2*0.10	2*P(t(100)>3.38199) <2*P(t(100)>3.174) < 2(0.001)
95% CI	$4.11 \pm (2.064)(1.43)/\sqrt{25}$ =(3.5, 4.7)	$4.11 \pm (1.96)(1.43)/\sqrt{1933}$ =(4.05, 4.17)

As n goes larger,

p-value goes smaller and confidence interval gets narrower.

9.52 Misleading summaries? Two researchers conduct separate studies to test H_0: $p = 0.50$ against H_a: $p \neq 0.50$, each with $n = 400$.

a. Researcher A gets 220 observations in the category of interest, and $\hat{p} = 220/400 = 0.550$ and test statistic $z = 2.00$. Show that the P-value $= 0.046$ for Researcher A's analysis.

b. Researcher B gets 219 in the category of interest, and $\hat{p} = 219/400 = 0.5475$ and test statistic $z = 1.90$. Show that the P-value $= 0.057$ for Researcher B's analysis.

c. Using $\alpha = 0.05$, indicate in each case from part a and part b whether the result is "statistically significant." Interpret.

d. From part a, part b, and part c, explain why important information is lost by reporting the result of a test as "P-value ≤ 0.05" versus "P-value > 0.05," or as "reject H_0" versus "do not reject H_0," instead of reporting the actual P-value.

e. Show that the 95% confidence interval for p is (0.501, 0.599) for Researcher A and (0.499, 0.596) for Researcher B. Explain how this method shows that, in practical terms, the two studies had very similar results.

	A	B
$\hat{p}$	220/400	219/400
z	2.00	1.90
p-value	0.046	0.057
Decision	Reject H_0	Do not reject H_0
Statistical significance	Yes	No
95% CI	(0.501, 0.599)	(0.499, 0.596)

Similar outcomes

Different conclusion

overlap a lot

c) Researcher A's result has a P-value less than 0.05; thus, it is "statistically significant." Researcher B's P-value is not less than 0.05, and is not, therefore, "statistically significant." Results that are not different from one another in practical terms might lead to different conclusions if based on statistical significance alone.

d) If we do not see these two P-values, but merely know that one is statistically significant and one is not, we are not able to see that the P-values are so similar.

e) We can add and subtract the result of (1.96)(0.025), that is the z-score associated with a 95% confidence interval multiplied by the standard error, to the mean for each of these samples to get confidence intervals of (0.501, 0.599) for Researcher A and (0.499, 0.596) for Researcher B. This method shows the enormous amount of overlap between the two confidence intervals. The plausible values for the population proportions are very similar in the two cases, which we would not realize by merely reporting whether the null was rejected in a test.

9.50 Detecting prostate cancer Refer to the previous exercise about medical diagnoses. A *New York Times* article (February 17, 1999) about the PSA blood test for detecting prostate cancer stated: "The test fails to detect prostate cancer in 1 in 4 men who have the disease."

a. For the PSA test, explain what a Type I error is, and explain the consequence to a man of this type of error.

b. For the PSA test, what is a Type II error? What is the consequence to a man of this type of error?

c. To which type of error does the probability of 1 in 4 refer?

d. The article also stated that if you receive a positive result, the probability that you do not actually have prostate cancer is 2/3. Explain the difference between this and the conditional probability of a Type I error, given that you do not actually have prostate cancer.

9.50

H_0: The patient does not have prostate cancer;

H_a: The patient does have prostate cancer.

Type I error occurs when we reject a correct H_0

Type II error occurs when we do not reject a false H_0

a) A Type I error would occur if we diagnose prostate cancer when there is none. This would be that a man would have treatment, or at least further testing, when he did not need any.

b) A Type II error would occur if we fail to diagnose prostate cancer when there is prostate cancer. This would mean that a man who had prostate cancer would not receive necessary treatment.

c) The probability of 1 in 4 refers to the probability of a Type II error.

d) The 2/3 refers to the probability that someone does not have prostate cancer given that he received a positive result. The probability of a Type I error, on the other hand, refers to the probability that someone will receive a positive result, given that he does not have prostate cancer.

9.60 Gender bias in selecting managers Exercise 9.19 tested the claim that female employees were passed over for management training in favor of their male colleagues. Statewide, the large pool of more than 1000 eligible employees who can be tapped for management training is 40% female and 60% male. Let p be the probability of selecting a female for any given selection. For testing H_0: $p = 0.40$ against H_a: $p < 0.40$ based on a random sample of 50 selections, using the 0.05 significance level, verify that:

a. A Type II error occurs if the sample proportion falls less than 1.645 standard errors below the null hypothesis value, which means that $\hat{p} > 0.286$.

b. When $p = 0.20$, a Type II error has probability 0.06.

a. Type II error occurs when we do not reject a wrong H_0.

Using the 0.05 significance level, do not reject H_0 means the p-value (left tail probability) > 0.05.

Which means do not reject H_0 when test statistic Z > -1645

$$z = (\hat{p} - p_0)/se_0 = (\hat{p} - 0.4)/(\sqrt{0.4 * 0.6/50}) > -1.645,$$

which means do not reject H_0 when $\hat{p} > 0.286$

b. When the true p=0.2,

the sampling distribution of $\hat{p}$ has mean p=0.2 and se=$\sqrt{(0.2)(1-0.2)/50}$=0.0566.

Then the probability of do not reject H_0 ($\hat{p} > 0.286$) is P(Z>(0.286-0.2)/0.0566) = 0.06.