Chapter 10

Hypothesis Tests Regarding a Parameter

A second type of statistical inference is hypothesis testing. Here, rather than use either a point (or interval) estimate from a simple random sample to approximate a population parameter, hypothesis testing uses point estimate to decide which of two hypotheses (guesses) about parameter is correct. We will look at hypothesis tests for proportion, p, mean, μ and standard deviation, σ.

10.1 The Language of Hypothesis Testing

In this section, we discuss hypothesis testing in general.

Exercise 10.1 (The Language of Hypothesis Testing)

1. *Test for binomial proportion, p, right-handed: defective batteries.*
   In a battery factory, 8% of all batteries made are assumed to be defective. Technical trouble with production line, however, has raised concern percent defective has increased in past few weeks. Of $n = 600$ batteries chosen at random, $\frac{70}{600}$ths $\left(\frac{70}{600} \approx 0.117\right)$ of them are found to be defective. Does data support concern about defective batteries at $\alpha = 0.05$?

   (a) *Statement.* Choose one.

      i. $H_0 : p = 0.08$ versus $H_1 : p < 0.08$
      ii. $H_0 : p \leq 0.08$ versus $H_1 : p > 0.08$
      iii. $H_0 : p = 0.08$ versus $H_1 : p > 0.08$

   (b) *Test.*

Chance $\hat{p} = \frac{70}{600} \approx 0.117$ or more, *if* $p_0 = 0.08$, is

$$\text{p–value} = P(\hat{p} \geq 0.117) = P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq \frac{0.117 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}}\right) \approx P\left(Z \geq 3.31\right) \approx$$

which equals (circle one) **0.00** / **0.04** / **4.65**.

(Stat, Proportions, One sample, with summary, Number of successes: 70, Number of observations: 600, Next, Null: prop. = 0.08 Alternative: > Calculate.)

Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(c) *Conclusion.*
Since p–value $= 0.0005 < \alpha = 0.05$,
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.08$.
So, sample $\hat{p}$ indicates population proportion p
is less than / **equals** / **is greater than** 0.08: $H_1 : p > 0.08$.

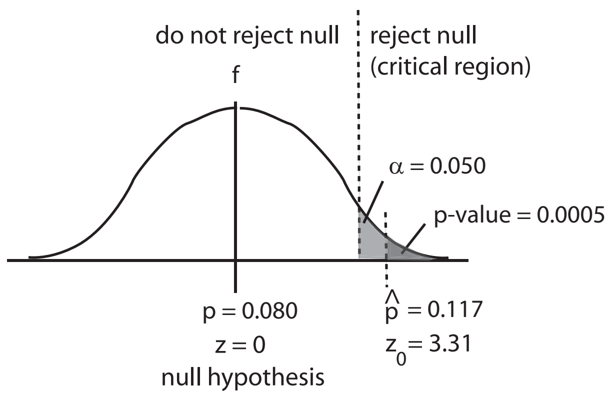

Figure 10.1 (P-value for sample $\hat{p} = 0.117$, if guess $p_0 = 0.08$)

(d) *A comment: null hypothesis and alternative hypothesis.*
In this hypothesis test, we are asked to choose between (choose one) **one** / **two** / **three** alternatives (or *hypotheses*, or guesses): a *null* hypothesis of $H_0 : p = 0.08$ and an *alternative* of $H_1 : p > 0.08$. *Null hypothesis is a statement of "status quo", of no change; test statistic used to reject it or not. Alternative hypothesis is "challenger" statement.*

(e) *Another comment: p-value.*
In this hypothesis test, p–value is chance observed proportion defective is 0.117 or more, guessing population proportion defective is $\boldsymbol{p_0 = 0.08}$ / $\boldsymbol{p_0 = 0.117}$.
In general, *p-value is probability of observing test statistic or more extreme, assuming null hypothesis is true.*

(f) *And another comment: test statistic different for different samples.*
If a second *sample* of 600 batteries were taken at *random* from all batteries, observed proportion defective of this second group of 600 batteries would

probably be (choose one) **different from / same as**
first observed proportion defective given above, $\hat{p} = 0.117$, say, $\hat{p} = 0.09$ which may change the conclusions of hypothesis test.

(g) *Possible mistake: Type I error.*
Even though hypothesis test tells us population proportion defective is greater than 0.08, that alternative $H_1 : p > 0.08$ is correct, we could be *wrong*. If we were indeed wrong, we should have picked **null / alternative** hypothesis $H_0 : p = 0.08$ instead. *Type I error is mistakenly rejecting null.* Also, $\alpha = P(\text{type I error})$.

2. *Test p, right–sided again: defective batteries.*
Of $n = 600$ batteries chosen at random, $\frac{54}{600}$ths $\left(\frac{54}{600} = 0.09\right)$, instead of 0.117, of them are found to be defective. Does data support concern about *increase* in defective batteries (from 0.08) at $\alpha = 0.05$ in this case?

   (a) *Statement.* Choose one.

      i. $H_0 : p = 0.08$ versus $H_1 : p < 0.08$
      ii. $H_0 : p \leq 0.08$ versus $H_1 : p > 0.08$
      iii. $H_0 : p = 0.08$ versus $H_1 : p > 0.08$

   (b) *Test.*
   Chance $\hat{p} = \frac{54}{600} = 0.09$ or more, *if* $p_0 = 0.08$, is

$$\text{p–value} = P(\hat{p} \geq 0.09) = P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq \frac{0.09 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}}\right) \approx P\left(Z \geq 0.903\right) \approx$$

   which equals (circle one) **0.00 / 0.09 / 0.18**.
   (Options, Number of successes: 54, Number of observations: 600, Next, Null: prop. = 0.08 Alternative: > Calculate.)
   Level of significance $\alpha =$ (choose one) **0.01 / 0.05 / 0.10**.

   (c) *Conclusion.*
   Since p–value $= 0.18 > \alpha = 0.05$,
   (circle one) **do not reject / reject** null guess: $H_0 : p = 0.08$.
   So, sample $\hat{p}$ indicates population proportion p
   is less than / equals / is greater than 0.08: $H_0 : p = 0.08$.

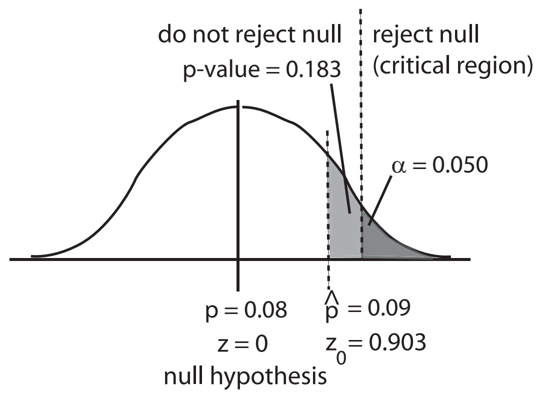

Figure 10.2 (P-value for sample $\hat{p} = 0.09$, if guess $p_0 = 0.08$)

(d) *Possible mistake: Type II error.*
Even though hypothesis test tells us population proportion defective is equal to null $H_0 : p = 0.08$, we could be *wrong*. If we were indeed wrong, we should have picked **null** / **alternative** hypothesis $H_1 : p > 0.08$. *Type II error is mistakenly rejecting alternative.* Also, $\beta = P(\text{type II error})$.

(e) *Comparing hypothesis tests.*
P–value associated with $\hat{p} = 0.09$ (p–value = 0.183) is
(choose one) **smaller** / **larger**
than p–value associated with $\hat{p} = 0.117$ (p–value = 0.00).
Sample proportion defective $\hat{p} = 0.09$ is
(choose one) **closer to** / **farther away from**,
than $\hat{p} = 0.117$, to null guess $p = 0.08$.
It makes sense we do not reject null guess of $p = 0.08$ when observed proportion is $\hat{p} = 0.09$, but reject null guess when observed proportion is $\hat{p} = 0.117$.

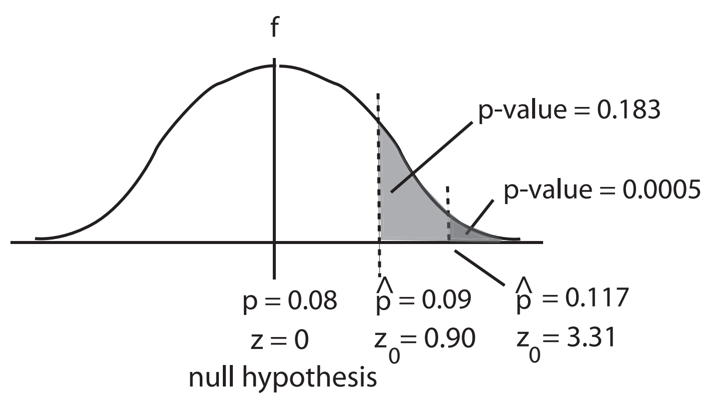

Figure 10.3 (Tend to reject null for smaller p-values)

If p–value is smaller than level of significance, α, reject null. If null is rejected when p–value is "really" small, less than $\alpha = 0.01$, test is *highly significant*. If null is rejected when p–value is small, typically between

$\alpha = 0.01$ and $\alpha = 0.05$, test is *significant*. If null is not rejected, test is *not significant*. This is one of *two* approaches to hypothesis testing.

(f) *Comparing Type I error (α) and Type II error (β).*

chosen ↓ actual →	null H_0 true	alternative H_1 true
choose null H_0	correct decision	type II error, β
choose alternative H_1	type I error, α	correct decision: power $= 1 - \beta$

As α decreases, β **decreases** / **increases** because if chance of mistakenly rejecting null decreases, this necessarily means chance of mistakenly rejecting alternative increases.

3. *Test p, left–sided this time: defective batteries.*
Of $n = 600$ batteries chosen at random, $\frac{36}{600}$ths $\left(\frac{36}{600} = 0.06\right)$ of them are found to be defective. Does data support possibility there is a *decrease* in defective batteries (from 0.08) at $\alpha = 0.05$ in this case?

(a) *Statement.* Choose one.

i. $H_0 : p = 0.08$ versus $H_1 : p < 0.08$
ii. $H_0 : p \leq 0.08$ versus $H_1 : p > 0.08$
iii. $H_0 : p = 0.08$ versus $H_1 : p > 0.08$

(b) *Test.*
Chance $\hat{p} = \frac{36}{600} = 0.06$ or *less*, *if* $p_0 = 0.08$, is

$$\text{p–value} = P(\hat{p} \leq 0.06) = P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \leq \frac{0.06 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}}\right) \approx P\left(Z \leq -1.81\right) \approx$$

which equals (circle one) **0.04** / **0.06** / **0.09**.

(Options, Number of successes: 36, Number of observations: 600, Next, Null: prop. = 0.08 Alternative: < (Notice: not >!) Calculate.)

Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(c) *Conclusion.*
Since p–value $= 0.04 < \alpha = 0.05$,
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.08$.
So, sample $\hat{p}$ indicates population proportion p
is less than / **equals** / **is greater than** 0.08: $H_1 : p < 0.08$.

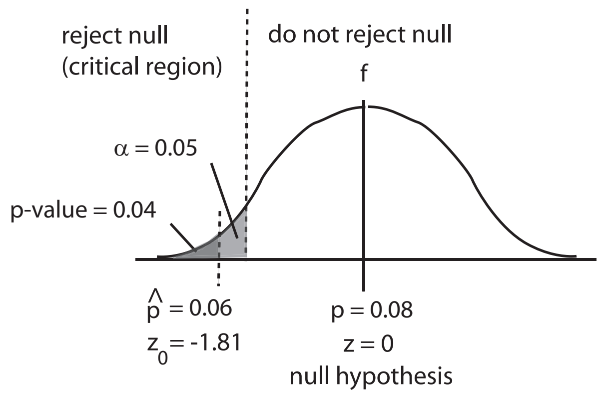

Figure 10.4 (P-value for sample $\hat{p} = 0.06$, if guess $p_0 = 0.08$)

(d) *One-sided tests.*
Right-sided test and left-sided tests are **one-sided** / **two-sided** tests.

4. *Test p, two-sided: defective batteries.*
Of $n = 600$ batteries chosen at random, $\frac{36}{600}$ths $\left(\frac{36}{600} = 0.06\right)$ of them are found to be defective. Does data support possibility there is a *change* in defective batteries (from 0.08) at $\alpha = 0.05$ in this case?

(a) *Statement.* Choose one.

i. $H_0 : p = 0.08$ versus $H_1 : p < 0.08$
ii. $H_0 : p \leq 0.08$ versus $H_1 : p > 0.08$
iii. $H_0 : p = 0.08$ versus $H_1 : p \neq 0.08$

(b) *Test.*
Chance $\hat{p} = \frac{36}{600} = 0.06$ or *less*, *or* $\hat{p} = 0.10$ or *more*, *if* $p_0 = 0.08$, is

$$\begin{aligned}
\text{p–value} &= P(\hat{p} \leq 0.06) + P(\hat{p} \geq 0.10) \\
&\approx P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \leq \frac{0.06 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}}\right) + P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq \frac{0.10 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}}\right) \\
&\approx P(Z \leq -1.81) + P(Z \geq 1.81)
\end{aligned}$$

which equals (circle one) **0.04** / **0.07** / **0.09**.

(Options, Number of successes: 36, Number of observations: 600, Next, Null: prop. = 0.08 Alternative: $\neq$ (Notice: not <!) Calculate.)

Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(c) *Conclusion.*
Since p–value $= 0.07 > \alpha = 0.05$,
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.08$.
So, sample $\hat{p}$ indicates population proportion p
is less than / **equals** / **is greater than** 0.08: $H_0 : p = 0.08$.

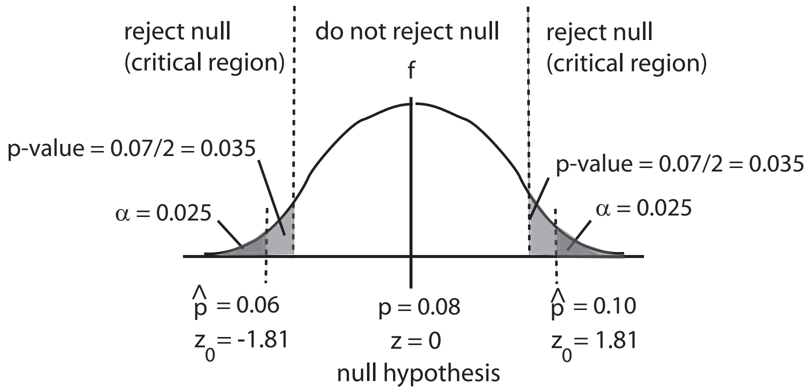

Figure 10.5 (Two-sided test)

(d) *One-sided test or two-sided test?*
If not sure whether population proportion p above or below null guess of $p_0 = 0.08$, use (choose one) **one-sided** / **two-sided** test.

(e) *Types of tests.*
In this course, we are interested in
right–sided test ($H_0 : p = p_0$ versus $H_1 : p > p_0$),
left–sided test ($H_0 : p = p_0$ versus $H_1 : p < p_0$) and
two–sided test ($H_0 : p = p_0$ versus $H_1 : p \neq p_0$).
However, other tests are possible; for example, (choose *one or more*!)

i. $H_0 : p < p_0$ versus $H_1 : p \geq p_0$
ii. $H_0 : p = p_0$ versus $H_1 : p = p_1$
iii. $H_0 : p_{0,1} \leq p \leq p_{0,2}$ versus $H_1 : p_{0,2} \leq p \leq p_{0,3}$

(f) *Steps in a test.*
Steps in any hypothesis test are (choose *one or more!*):

i. Statement
ii. Test
iii. Conclusions

10.2 Hypothesis Tests for a Population Proportion

Hypothesis test for p from binomial has test statistic

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

where we assume a large simple random sample has been chosen, $np_0(1-p_0) \geq 10$. If sampled from finite population, $n \leq 0.05N$. We consider three approaches to testing: p-value, classical and confidence interval. We will also discuss one-sided testing for p

when sample size is small.

Although all three approaches are mentioned, emphasis will be on p-value approach throughout course.

Exercise 10.2 (Hypothesis Tests for a Population Proportion)

1. *Test p: defective batteries.*
   Of $n = 600$ batteries chosen at random, $\frac{54}{600}$ths $\left(\frac{54}{600} = 0.09\right)$ of them are found to be defective. Does data support hypotheses of an *increase* in defective batteries (from 0.08) at $\alpha = 0.05$ in this case? Solve using both p-value and classical approaches to hypothesis testing.

   (a) *Check assumptions.*
   Since $np_0(1 - p_0) = 600(0.08)(1 - 0.08) = 44.16 > 10$, assumptions **have** / **have not** been satisfied, so continue large sample test.

   (b) *P-value approach (review).*

   i. *Statement.* Choose one.
   A. $H_0 : p = 0.08$ versus $H_1 : p < 0.08$
   B. $H_0 : p \leq 0.08$ versus $H_1 : p > 0.08$
   C. $H_0 : p = 0.08$ versus $H_1 : p > 0.08$

   ii. *Test.*
   Chance $\hat{p} = \frac{54}{600} = 0.09$ or more, *if* $p_0 = 0.08$, is

   $$\text{p–value} = P(\hat{p} \geq 0.09) = P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq \frac{0.09 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}}\right) \approx P\left(Z \geq 0.903\right) \approx$$

   which equals (circle one) **0.00** / **0.09** / **0.18**.
   (Stat, Proportions, One sample, with summary, Number of successes: 54, Number of observations: 600, Next, Null: prop. = 0.08 Alternative: > Calculate.)
   Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

   iii. *Conclusion.*
   Since p–value $= 0.18 > \alpha = 0.05$,
   (circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.08$.
   So, sample $\hat{p}$ indicates population proportion p
   is less than / **equals** / **is greater than** 0.08: $H_0 : p = 0.08$.

   (c) *Classical approach.*

   i. *Statement.* Choose one.
   A. $H_0 : p = 0.08$ versus $H_1 : p < 0.08$
   B. $H_0 : p \leq 0.08$ versus $H_1 : p > 0.08$
   C. $H_0 : p = 0.08$ versus $H_1 : p > 0.08$

ii. *Test.*
Test statistic
$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx \frac{0.09 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}} \approx$$
0.90 / **1.99** / **2.18**.
Critical value at $\alpha = 0.05$,
$$z_\alpha = z_{0.05} \approx$$
1.28 / **1.65** / **2.58**
(Stat, Calculators, Normal, Mean: 0, Std. Dev.: 1, Prob(X ≥ $\boxed{?}$) = 0.05 Calculate.)

iii. *Conclusion.*
Since $z_0 = 0.903 < z_{0.05} = 1.645$,
(test statistic is outside critical region, so "close" to null 0.08),
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.08$.
So, sample $\hat{p}$ indicates population proportion p
is less than / **equals** / **is greater than** 0.08: $H_0 : p = 0.08$.

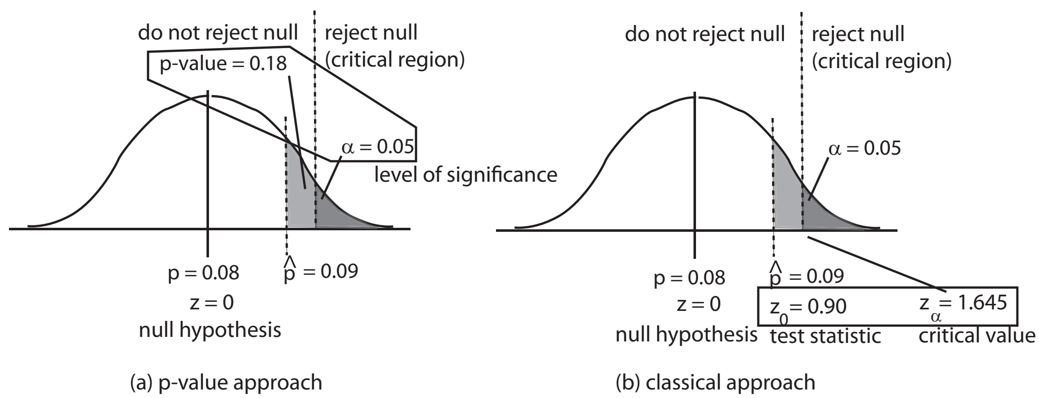

Figure 10.6 (P-value versus Classical Approach)

(d) **True** / **False**
When we *do not reject* null, we pick hypothesis with "equals" in it.

(e) To say we *do not reject* null means, in this case, (circle *one or more*)

i. disagree with claim percent defective of batteries is greater than 8%.

ii. data does not support claim percent defective of batteries is greater than 8%.

iii. we fail to reject percent defective of batteries equals 8%.

(f) *Population, Sample, Statistic, Parameter.* Match columns.

terms	battery example
(a) population	**(a)** all (defective or nondefective) batteries
(b) sample	**(b)** proportion defective, of all batteries, p
(c) statistic	**(c)** 600 (defective or nondefective) batteries
(d) parameter	**(d)** proportion defective, of 600 batteries, $\hat{p}$

terms	(a)	(b)	(c)	(d)
example				

2. *Test for binomial proportion, p: overweight in Indiana.*
An investigator wishes to know whether proportion of overweight individuals in Indiana *differs* from national proportion of 71% or not. A random sample of size $n = 600$ results in 450 $\left(\frac{450}{600} = 0.75\right)$ who are overweight. Test at $\alpha = 0.05$.

(a) *Check assumptions.*
Since $np_0(1 - p_0) = 600(0.71)(1 - 0.71) = 123.54 > 10$, assumptions **have** / **have not** been satisfied, so continue large sample test.

(b) *P-value approach.*

i. *Statement.* Choose one.

A. $H_0 : p = 0.71$ versus $H_1 : p > 0.71$

B. $H_0 : p = 0.71$ versus $H_1 : p < 0.71$

C. $H_0 : p = 0.71$ versus $H_1 : p \neq 0.71$

ii. *Test.*
Since test statistic of $\hat{p} = \frac{450}{600} = 0.75$ is

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.75 - 0.71}{\sqrt{\frac{0.71(1-0.71)}{600}}} =$$

(circle one) **1.42** / **1.93** / **2.16**,
chance $\hat{p} = -0.75$ or less or $\hat{p} = 0.75$ or more, *if* $p_0 = 0.71$, is

$$\text{p-value } = P(Z \leq -2.16) + P(Z \geq 2.16) = 2 \times P(X \geq 2.16) \approx$$

(choose closest one) **0.031** / **0.057** / **0.075**.

(Stat, Proportions, One sample, with summary, Number of successes: 450, Number of observations: 600, Next, Null: prop. = 0.71 Alternative: $\neq$ Calculate.)

Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

iii. *Conclusion.*
Since p–value $= 0.031 < \alpha = 0.050$,
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.71$.
In other words, sample $\hat{p}$ indicates population proportion p
(circle one) **equals** / **does not equal** 0.71: $H_1 : p \neq 0.71$.

(c) *Classical approach.*

i. *Statement.* Choose one.

A. $H_0 : p = 0.71$ versus $H_1 : p > 0.71$
B. $H_0 : p = 0.71$ versus $H_1 : p < 0.71$
C. $H_0 : p = 0.71$ versus $H_1 : p \neq 0.71$

ii. *Test.*
Test statistic

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx \frac{0.75 - 0.71}{\sqrt{\frac{0.71(1-0.71)}{600}}} \approx$$

0.90 / **1.99** / **2.16**.
Critical values at $\pm\alpha = 0.05$,

$$\pm z_{\frac{\alpha}{2}} = \pm z_{\frac{0.05}{2}} = \pm z_{0.025} \approx$$

±1.28 / **±1.65** / **±1.96**

(Stat, Calculators, Normal, Mean: 0, Std. Dev.: 1, Prob(X ≥ $\boxed{?}$) = 0.025 Calculate.)

iii. *Conclusion.*
Since $-z_{0.025} = -1.96 < z_{0.025} = 1.96 < z_0 = 2.16$
(test statistic is inside critical region, so "far away" from null 0.71),
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.71$.
In other words, sample $\hat{p}$ indicates population proportion p
(circle one) **equals** / **does not equal** 0.71: $H_1 : p \neq 0.71$.

(d) *95% Confidence Interval (CI) for p.*
The 95% CI for proportion of overweight individuals, p, is
(0.018, 0.715) / **(0.715, 0.785)** / **(0.533, 0.867)**.
(Stat, Proportions, one sample, with summary, Number of successes: 450, Number of observations: 600, choose Confidence Interval 0.95, Calculate.)
Since null $p_0 = 0.71$ is *out*side CI (0.715, 0.785),
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.71$.
In other words, sample $\hat{p}$ indicates population proportion p
(circle one) **equals** / **does not equal** 0.71: $H_1 : p \neq 0.71$.

3. *Test for binomial proportion, p, small sample: conspiring Earthlings.*
It appears 6.5% of Earthlings are conspiring with little green men (LGM) to take over Earth. Human versus Extraterrestrial Legion Pact (HELP) claims more than 6.5% of Earthlings are conspiring with LGM. In a random sample of 100 Earthlings, 7 $\left(\frac{7}{100} = 0.07\right)$ are found to be conspiring with little green men (LGM). Does this data support HELP claim at $\alpha = 0.05$?

(a) *Check assumptions.*
Since $np_0(1 - p_0) = 100(0.065)(1 - 0.065) \approx 6.1 < 10$, assumptions **have** / **have not** been satisfied, so try *small* sample test instead.

(b) *Statement.* Choose one.

i. $H_0 : p = 0.065$ versus $H_1 : p < 0.065$
ii. $H_0 : p \leq 0.065$ versus $H_1 : p > 0.065$
iii. $H_0 : p = 0.065$ versus $H_1 : p > 0.065$

(c) *Test.*
Chance $\hat{p} = \frac{7}{100} = 0.070$ or more, *if* $p_0 = 0.065$, is

$$\text{p–value} = P(\hat{p} \geq 0.07) = P(X \geq 7) \approx$$

(choose closest one) **0.22** / **0.32** / **0.48**.
(Stat, Calculators, Binomial, n: 100, p: 0.065, Prob(X $\geq$ 7) = $\boxed{?}$ Calculate.)
Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(d) *Conclusion.*
Since p–value $= 0.48 > \alpha = 0.05$,
(circle one) **do not reject** / **reject** null guess: $H_0 : p = 0.065$.
So, sample $\hat{p}$ indicates population proportion p
is less than / **equals** / **is greater than** 0.065: $H_0 : p = 0.065$.

(e) HELP claim in this test is there are more than 6.5% conspiring Earthlings. Claim in this test, *any* test, for that matter, is *always* a statement about (circle one) **null** / **alternative** hypothesis.

(f) Test constructed so if there is any doubt as to validity of HELP claim, we will fall back on not rejecting null, there are 6.5% conspiring Earthlings. Hypothesis test (always) favors (circle one) **null** / **alternative** hypothesis.

(g) The p–value, in this case, is chance (choose one)

i. population proportion 0.07 or more, if observed proportion 0.065.
ii. observed proportion 0.07 or more, if observed proportion 0.065.
iii. population proportion 0.07 or more, if population proportion 0.065.
iv. observed proportion 0.07 or more, if population proportion 0.065.

10.3 Hypothesis Tests for a Population Mean

Hypothesis test for μ with *un*known σ is a *t-test* with test statistic

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}},$$

and used when either underlying distribution is normal with no outliers or if simple random sample size large ($n \geq 30$).

Exercise 10.3 (Mean–Population Standard Deviation Unknown)

1. *Testing μ, right–sided: hourly wages.*
Average hourly wage in US is assumed to be \$10.05 in 1985. Midwest big business, however, claims average hourly wage to be larger than this. A random sample of size $n = 15$ of midwest workers determines average hourly wage $\bar{x} = \$10.83$ and standard deviation in wages $s = 3.25$. Does data support big business's claim at $\alpha = 0.05$? Assume normality.

   (a) *Statement.* Choose one.

      i. $H_0 : \mu = \$10.05$ versus $H_1 : \mu < \$10.05$
      ii. $H_0 : \mu = \$10.05$ versus $H_1 : \mu \neq \$10.05$
      iii. $H_0 : \mu = \$10.05$ versus $H_1 : \mu > \$10.05$

   (b) *Test.*
   Chance $\bar{x} = 10.83$ or more, *if* $\mu_0 = 10.05$, is

$$\text{p–value} = P(\bar{X} \geq 10.83) = P\left(\frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \geq \frac{10.83 - 10.05}{\frac{3.25}{\sqrt{15}}}\right) \approx P\left(t \geq 0.93\right) \approx$$

   which equals (circle one) **0.18** / **0.20** / **0.23**.
   (Stat, T Statistics, with summary, Sample mean: 10.83, Sample std. dev.: 3.25, Sample size: 15, Next, Null: mean = 10.05, Alternative: > Calculate.)
   Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

   (c) *Conclusion.*
   Since p–value $= 0.18 > \alpha = 0.05$,
   (circle one) **do not reject** / **reject** null guess: $H_0 : \mu = 10.05$.
   In other words, sample $\bar{x}$ indicates population average salary μ
   is less than / **equals** / **is greater than** 10.05: $H_0 : \mu = 10.05$.

   (d) Match columns.

terms	wage example
(a) population	**(a)** all US wages
(b) sample	**(b)** wages of 15 workers
(c) statistic	**(c)** observed average wage of 15 workers
(d) parameter	**(d)** average wage of US workers, μ

terms	(a)	(b)	(c)	(d)
wage example				

2. *Testing μ, right–sided: accounting program.*
Program director for an accounting program wishes to test, at 5%, hypothesis her students score higher than national average of 615 on national final exam. She randomly selects 11 recent graduates of two–year program and discovers $\bar{x} = 630$, and $s = 23$. Assume underlying distribution is normal.

(a) *Statement.* Choose one.

i. $H_0 : \mu = 615$ versus $H_1 : \mu < 615$
ii. $H_0 : \mu = 615$ versus $H_1 : \mu \neq 615$
iii. $H_0 : \mu = 615$ versus $H_1 : \mu > 615$

(b) *Test (a slightly different way).*
Since test statistic of $\bar{x} = 630$ is

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{630 - 615}{23/\sqrt{11}} =$$

(circle one) **0.42** / **0.93** / **2.16**,
chance $\bar{x} = 630$ or more, *if* $\mu_0 = 615$, is
p–value $= P(t \geq 2.16) \approx$ (circle one) **0.03** / **0.20** / **0.23**.
(Stat, T Statistics, with summary, Sample mean: 630, Sample std. dev.: 23, Sample size: 11, Next, Null: mean = 615, Alternative: > Calculate.)
Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(c) *Conclusion.*
Since p–value $= 0.03 < \alpha = 0.05$,
(circle one) **do not reject** / **reject** null guess: $H_0 : \mu = 615$.
In other words, sample $\bar{x}$ indicates program population average score μ **is less than** / **equals** / **is greater than** 615: $H_1 : \mu > 615$.

(d) *A comment.*
True / **False** To say that this data indicates true (population) average score of *her* students score higher than national average also means true (population) average of *all* students in nation score higher on national average.

(e) *Related question: critical value.*
The critical value is $t_\alpha = t_{0.05} \approx$ (choose one) **−1.81** / **0.81** / **1.81**.
(Stat, Calculators, T, Mean: 10, Prob(X $\geq \boxed{?}$) = 0.05 Calculate.)

3. *Testing μ, right–sided, raw data: sprinkler activation time.*
Thirteen data values are observed in a fire-prevention study of sprinkler activation times (in seconds).

27 41 22 27 23 35 30 33 24 27 28 22 24

Actual average activation time is supposed to be 25 seconds. Test if it is more than this at 5%.
(StatCrunch: Blank data table. Relabel var1 as time. Type 13 lengths into time column. Data, Save data, 10.3 sprinkler times. OK)

(a) *Check assumptions (since $n = 13 < 30$): normality and outliers.*

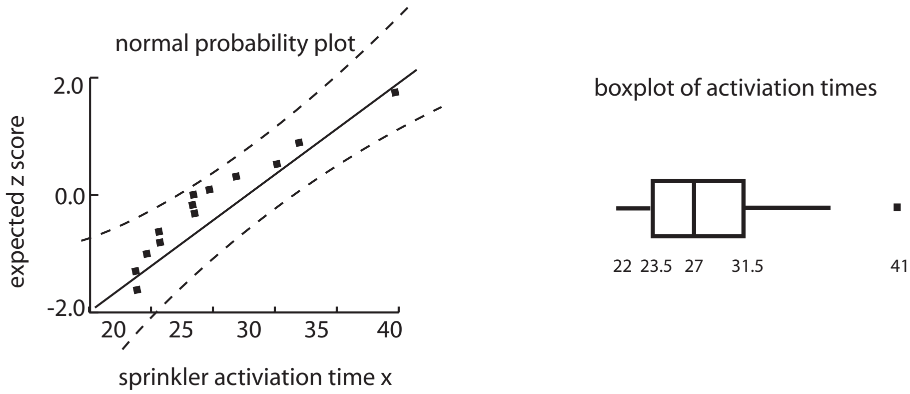

Figure 10.7 (Normal probability plot, boxplot for times.)

i. *Data normal?*
Normal probability plot indicates activation times
normal / **not normal** because data within dotted bounds.
Graphics. QQ Plot, Select Columns: time, Create Graph!

ii. *Outliers?*
Boxplot indicates **outliers** / **no outliers** (but ignore it).
Graphics. Boxplot, select time, Next, Use fences to identify outliers, check Draw boxes horizontally. Create Graph!

(b) *Statement.* Choose one.

i. $H_0 : \mu = 25$ versus $H_1 : \mu > 25$
ii. $H_0 : \mu = 25$ versus $H_1 : \mu < 25$
iii. $H_0 : \mu = 25$ versus $H_1 : \mu \neq 25$

(c) *Test.*
Chance $\bar{x} = 27.9$ or more, *if* $\mu_0 = 25$, is

$$\text{p–value} = P(\bar{X} \geq 27.9) = P\left(\frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \geq \frac{27.9 - 25}{\frac{5.6}{\sqrt{13}}}\right) \approx P\left(t \geq 1.88\right) \approx$$

(choose one) **0.039** / **0.043** / **0.341**.
(Stat, T Statistics, one sample, with data, Next, Null: mean = 25, Alternative: > Calculate.)
Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(d) *Conclusion.*
Since p–value $= 0.043 < \alpha = 0.050$,
(circle one) **do not reject** / **reject** null guess: $H_0 : \mu = 25$.
In other words, sample $\bar{x}$ indicates population average time μ
is less than / **equals** / **is greater than** 25: $H_0 : \mu > 25$.

(e) *Related question: unknown σ.*
The σ is (choose one) **known** / **unknown** in this case
and is approximated by $s \approx$ (choose one) **5.62** / **5.73** / **5.81**.
(Stat, Summary Stats, Select Columns: time, Calculate.)

(f) *Related question: critical value.*
The critical value is $t_\alpha = t_{0.05} \approx$ (choose one) **−1.78** / **−0.78** / **1.78**.
(Stat, Calculators, T, DF: 12, Prob(X $\geq \boxed{?}$) = 0.05 Calculate.)

4. *Testing μ, left–sided: weight of coffee*
Label on a large can of Hilltop Coffee states average weight of coffee contained in all cans it produces is 3 pounds of coffee. A coffee drinker association claims average weight is less than 3 pounds of coffee, $\mu < 3$. Suppose a random sample of 30 cans has an average weight of $\bar{x} = 2.95$ pounds and standard deviation of $s = 0.18$. Does data support coffee drinker association's claim at $\alpha = 0.05$?

(a) *P-value approach.*

i. *Statement.* Choose one.
A. $H_0 : \mu = 3$ versus $H_1 : \mu < 3$
B. $H_0 : \mu < 3$ versus $H_1 : \mu > 3$
C. $H_0 : \mu = 3$ versus $H_1 : \mu \neq 3$

ii. *Test.*
Chance observed $\bar{x} = 2.95$ or *less*, *if* $\mu_0 = 3$, is

$$\text{p–value} = P(\bar{X} \leq 2.95) = P\left(\frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \leq \frac{2.95 - 3}{\frac{0.18}{\sqrt{30}}}\right) \approx P\,(t \leq -1.52)$$

which equals (choose one) **0.04** / **0.07** / **0.08**.
(Stat, T Statistics, with summary, Sample mean: 2.95, Sample std. dev.: 0.18, Sample size: 30, Next, Null: mean = 3, Alternative: < Calculate.)
Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

iii. *Conclusion.*
Since p–value $= 0.07 > \alpha = 0.05$,
(circle one) **do not reject** / **reject** the null guess: $H_0 : \mu = 3$.
In other words, sample $\bar{x}$ indicates population average weight μ
(circle one) **is less than** / **equals** / **is greater than** 3: $H_0 : \mu = 3$.

(b) *Classical approach.*

i. *Statement.* Choose one.
A. $H_0 : \mu = 3$ versus $H_1 : \mu < 3$
B. $H_0 : \mu < 3$ versus $H_1 : \mu > 3$
C. $H_0 : \mu = 3$ versus $H_1 : \mu \neq 3$

ii. *Test.*
Test statistic

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{2.95 - 3}{0.18/\sqrt{30}} \approx$$

(choose one) $\mathbf{-1.65}$ / $\mathbf{-1.52}$ / $\mathbf{-1.38}$.
Critical value at $\alpha = 0.05$,

$$t_\alpha = -t_{0.05} \approx$$

(choose one) $\mathbf{-1.14}$ / $\mathbf{-1.70}$ / $\mathbf{-2.34}$.
(Stat, Calculators, T, DF: 29, Prob(X $\leq \boxed{?}$) = 0.05 Calculate.)

iii. *Conclusion.*
Since $t_0 \approx -1.52 > -t_{0.05} \approx -1.70$,
(test statistic is outside critical region, so "close" to null $\mu = 3$),
(circle one) **do not reject** / **reject** the null guess: $H_0 : \mu = 3$.
In other words, sample $\bar{x}$ indicates population average weight μ
(circle one) **is less than** / **equals** / **is greater than** 3: $H_0 : \mu = 3$.

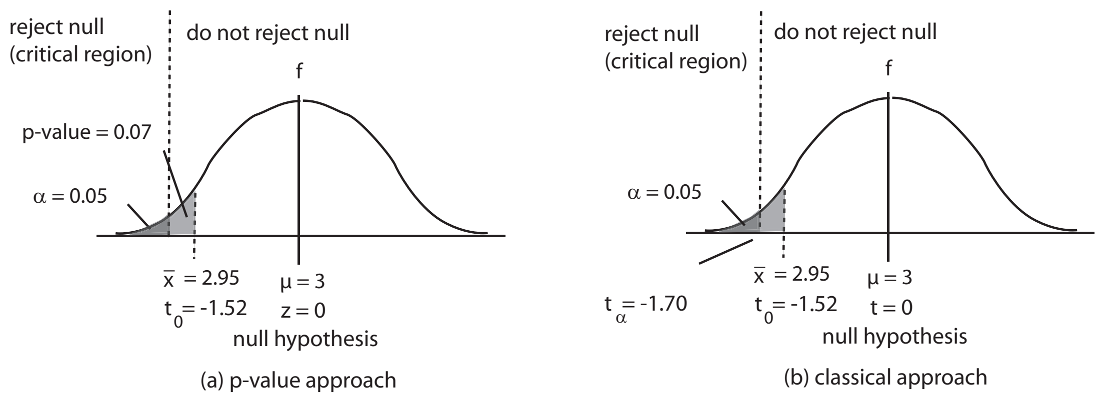

Figure 10.8 (Left-sided test of μ: coffee)

(c) *Related questions.*

i. "Claim" in this test, *any* test, is *always* **null** / **alternative** hypothesis.

ii. Test *always* favors **null** / **alternative** hypothesis since chance of mistakenly rejecting null is so small; in this case, $\alpha = 0.05$.

iii. To assume null hypothesis true, means assuming (circle one)

A. average weight of coffee contained in all Hilltop Coffee cans is less than 3 pounds, $\mu < 3$.

B. average weight of coffee contained in all Hilltop Coffee cans equals 3 pounds, $\mu = 3$.

C. average weight of coffee contained in a random sample of 30 Hilltop Coffee cans equals 2.95 pounds, $\bar{x} = 2.95$.

iv. Match columns.

terms	coffee example
(a) population	**(a)** average weight of all cans
(b) sample	**(b)** average weight of 30 cans
(c) statistic	**(c)** weights of 30 cans
(d) parameter	**(d)** weights of all cans

terms	(a)	(b)	(c)	(d)
coffee example				

10.4 Hypothesis Tests for a Population Standard Deviation

Hypothesis test for σ is a χ^2-*test* with test statistic

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2},$$

and used when either underlying distribution is normal with no outliers and sample is obtained using simple random sampling[1].

Exercise 10.4 (Hypothesis Tests for a Population Standard Deviation)

1. *Test for σ: car door and jamb.*
   In a simple random sample of 28 cars, SD in gap between door and jamb is $s = 0.7$ mm. Test if SD is *greater* than 0.6 mm at $\alpha = 0.05$. Assume normality with no outliers.

   (a) *Statement.* Choose one.

   i. $H_0 : \sigma = 0.6$ versus $H_1 : \sigma > 0.6$
   ii. $H_0 : \sigma = 0.6$ versus $H_1 : \sigma < 0.6$
   iii. $H_0 : \sigma = 0.6$ versus $H_1 : \sigma \neq 0.6$

   (b) *Test (p-value approach).*
   Since degrees of freedom

   $$\text{df} = n - 1 = 28 - 1 =$$

   (circle one) **15** / **27** / **34**,
   chance $s = 0.7$ or more, *if* $\sigma = 0.6$, is

   $$\text{p–value} = P(s \geq 0.7) = P\left(\frac{(n-1)s^2}{\sigma^2} \geq \frac{(28-1)(0.7^2)}{0.6^2}\right) \approx P\left(\chi^2 \geq 36.75\right) \approx$$

[1] This test is used not only for SD σ but also variance σ^2. CI is sensitive to non-normal data which is not always fixed by large sample size.

(circle one) **0.01** / **0.05** / **0.10**.

(Stat, Variance, with summary, Sample variance: 0.49, Sample size: 28, Next, Null: mean = 0.36, Alternative: > Calculate. Notice, SDs must be squared to variances to use StatCrunch.)

Level of significance $\alpha =$ (choose one) **0.01** / **0.05** / **0.10**.

(c) *Conclusion.*
Since p–value $= 0.10 > \alpha = 0.05$,
(circle one) **do not reject** / **reject** null guess: $H_0 : \sigma = 0.6$.
So, sample s indicates population SD σ
is less than / **equals** / **is greater than** 0.6: $H_0 : \sigma = 0.6$.

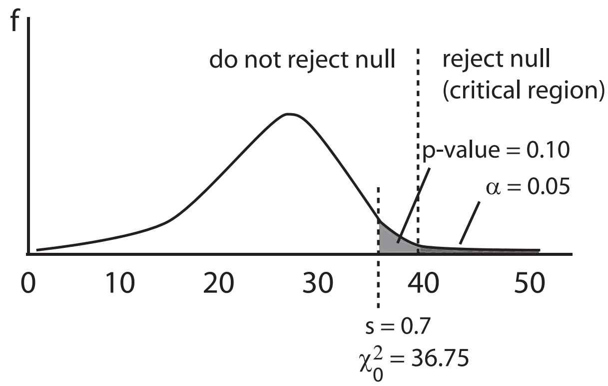

Figure 10.9 (Right-sided χ^2 test of σ)

(d) *Population, parameter, sample and statistic.* Match columns.

terms	jamb example
(a) population	**(a)** SD in jamb–door distance, of 28 cars, s
(b) sample	**(b)** SD in jamb–door distance, of all cars, σ
(c) statistic	**(c)** jamb–door distances, of all cars
(d) parameter	**(d)** jamb–door distances, of 28 cars

terms	(a)	(b)	(c)	(d)
jamb example				

2. *Inference for variance: machine parts.*
In a random sample of 18 machine parts, SD in lengths is $s = 12$ mm. Test if SD is *less* than 13 mm at $\alpha = 0.05$.

(a) *Statement.* The statement of the test is (circle one)

i. $H_0 : \sigma = 13$ versus $H_1 : \sigma > 13$
ii. $H_0 : \sigma = 13$ versus $H_1 : \sigma < 13$
iii. $H_0 : \sigma = 13$ versus $H_1 : \sigma \neq 13$

(b) *Test (classical approach).*
Test statistic of $s = 12$ is

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(18-1)(12)^2}{13^2} =$$

(circle one) **14.48** / **60.41** / **82.47**.
(Stat, Variance, with summary, Sample variance: 144, Sample size: 18, Next, Null: mean = 169, Alternative: < Calculate. Again notice SDs must be squared to variances to use StatCrunch.)
with degrees of freedom

$$\text{n} - 1 = 18 - 1 =$$

(circle one) **15** / **17** / **34** df,
Lower critical value at $\alpha = 0.05$,

$$\chi^2_{1-\alpha} = \chi^2_{0.95} \approx$$

(choose one) **7.67** / **8.67** / **9.67**.
(Stat, Calculators, Chi-square, DF: 17, Prob(X $\geq \boxed{?}$) = 0.95 Calculate.)

(c) *Conclusion.*
Since $\chi_0^2 \approx 14.45 > \chi^2_{0.95} \approx 8.67$,
(test statistic is outside critical region, so "close" to null $\sigma = 13$),
(circle one) **do not reject** / **reject** null guess: $H_0 : \sigma = 13$.
In other words, sample s indicates population SD σ
(circle one) **is less than** / **equals** / **is greater than** 13: $H_0 : \sigma = 13$.

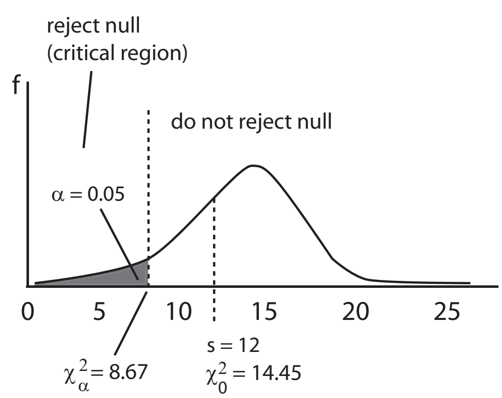

Figure 10.10 (Left-sided χ^2 test of σ)

10.5 Putting It Together: Which Method Do I Use?

The following table summaries all of the hypothesis tests given in this chapter and under what circumstances to calculate any one of these hypothesis tests; other hypothesis tests are given in later chapters. This table has also been used in earlier chapters to summarize confidence intervals.

HYPOTHESIS TESTS		mean μ	variance σ^2	proportion p
	one	$z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$	large: $z_0 = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ small: use binomial
sample	two	chapter 11	chapter 11	chapter 11
	multiple	chapter 13	not covered	chapter 12

10.6 The Probability of a Type II Error and the Power of the Test

Not covered.