Chapter 5

Hypothesis Testing

A second type of statistical inference is hypothesis testing. Here, rather than use either a point (or interval) estimate from a 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, and mean, μ, and standard deviation, σ.

5.1 Hypothesis Testing

In this section, we discuss hypothesis testing in general.

Exercise 5.1 (Introduction)

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. $H_0 : p \le 0.08$ 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

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

which equals (i) **0.00** (ii) **0.04** (iii) **4.65**.

```
prop1.test <- function(x, n, p.null, signif.level, type) {
  p.hat <- x/n
  z.test.statistic <- (p.hat-p.null)/(sqrt(p.null*(1-p.null)/n))
  if(type=="right") {
    z.crit <- -1*qnorm(signif.level)
    p.value <- 1-pnorm(z.test.statistic)
  }
  if(type=="left") {
    z.crit <- qnorm(signif.level)
    p.value <- pnorm(z.test.statistic)
  }
  if(type=="two.sided") {
    z.crit <- c(qnorm(signif.level/2),-1*qnorm(signif.level/2))
    p.value <- 2*min(1-pnorm(z.test.statistic),pnorm(z.test.statistic))
  }
  dat <- c(p.null, p.hat, z.crit, z.test.statistic, p.value)
  if(type=="two.sided") names(dat) <- c("p.null", "p.hat", "lower z crit", "upper z crit", "z test stat", "p value"
  if(type != "two.sided") names(dat) <- c("p.null", "p.hat", "z crit Value", "z test stat", "p value")
  return(dat)
}
prop1.test(70, 600, 0.08, 0.05, "right") # approx 1-proportion test for p

      p.null         p.hat z crit Value  z test stat       p value
0.0800000000 0.1166666667 1.6448536270 3.3106109599 0.0004654627
```

Level of significance $\alpha =$ (i) **0.01** (ii) **0.05** (iii) **0.10**.

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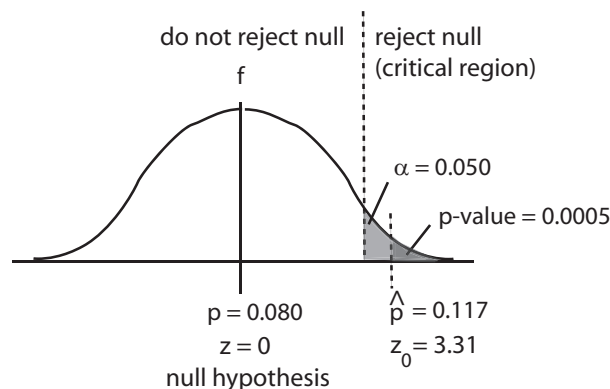

Figure 5.1: P-value for statistic $\hat{p} = 0.117$, if guess $p_0 = 0.08$

(d) *A comment: null hypothesis and alternative hypothesis.*
   In this hypothesis test, we are asked to choose between
   (i) **one** (ii) **two** (iii) **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 (i) $p_0 = 0.08$ (ii) $p_0 = 0.117$. *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 (i) **different from** (ii) **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 (i) **null** (ii) **alternative** hypothesis $H_0 : p = 0.08$ instead. *Type I error is mistakenly rejecting null; also, $\alpha = P(type\ I\ error)$.*

(h) *Binomial probability model.*
Let random variable X represent the number of $n = 600$ batteries which are defective and let parameter p be the probability a battery is defective and so an appropriate model for X is binomial pmf,

$$f(x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x}, \ x = 0, 1, \ldots, n,$$

where the null hypothesis is (i) $p = 0.08$ (ii) $p > 0.08$,
and alternative hypothesis is (i) $p = 0.08$ (ii) $p > 0.08$.

(i) *Normal approximation to binomial model.*
Let $\hat{p} = \frac{x}{n} = \frac{54}{600}$ estimate parameter p. Because of the central limit theorem, the probability distribution of statistic $\hat{P} = \frac{X}{n}$ can be approximated by the normal distribution with mean $\mu = p$ and $\sigma = \sqrt{\frac{p(1-p)}{n}}$ with z-score

$$Z = \frac{\hat{P} - \mu}{\sigma} = \frac{\hat{P} - p}{\sqrt{\frac{p(1-p)}{n}}}.$$

where Z is also normal but mean $\mu = 0$ and $\sigma = 1$.
(i) **True** (ii) **False**

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 (i) **0.00** (ii)**0.09** (iii) **0.18**.

   ```
   prop1.test(54, 600, 0.08, 0.05, "right") # approx 1-proportion test for p, right-sided

        p.null        p.hat z crit value  z test stat     p value
     0.0800000    0.0900000    1.6448536    0.9028939   0.1832911
   ```

   Level of significance $\alpha =$ (i) **0.01** (ii) **0.05** (iii) **0.10**.

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

   (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 (i) **null** (ii) **alternative** hypothesis $H_1 : p > 0.08$. *Type II error is mistakenly rejecting alternative; also, $\beta = P(type II error)$.*

   (e) *Comparing hypothesis tests.*
   P–value associated with $\hat{p} = 0.09$ (p–value $= 0.183$) is
   (i) **smaller** (ii) **larger**
   than p–value associated with $\hat{p} = 0.117$ (p–value $= 0.00$).
   Sample proportion defective $\hat{p} = 0.09$ is
   (i) **closer to** (ii) **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
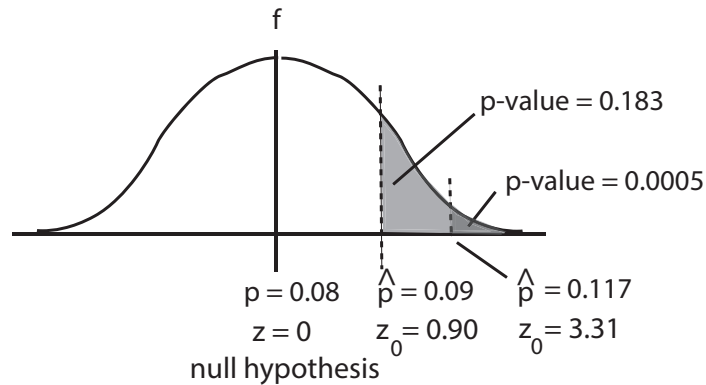
Figure 5.2: Tend to reject null for smaller p-values

$\hat{p} = 0.117$. *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.

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 the claim there is a *decrease* in defective batteries (from 0.08) at $\alpha = 0.05$ in this case?
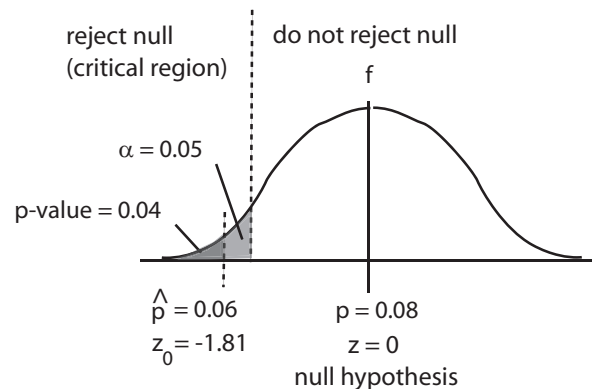
Figure 5.3: P-value for statistic $\hat{p} = 0.06$, if guess $p_0 = 0.08$

(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

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

which equals (i) **0.04** (ii) **0.06** (iii) **0.09**.

```
prop1.test(36, 600, 0.08, 0.05, "left") # approx 1-proportion test for p, left-sided

      p.null        p.hat z crit value  z test stat     p value
   0.08000000   0.06000000  -1.64485363  -1.80578780  0.03547575
```

Level of significance $\alpha =$ (i) **0.01** (ii) **0.05** (iii) **0.10**.

(c) *Conclusion.*
(Technical.) Since p–value $= 0.04 < \alpha = 0.05$,
(i) **do not reject** (ii) **reject** null guess: $H_0 : p = 0.08$.
(Final.) So, sample $\hat{p}$ indicates population proportion p
(i) **is less than** (ii) **equals** (iii) **is greater than** 0.08: $H_1 : p < 0.08$.
Statistic $\hat{p} = 0.06$ (i) **proves** (ii) **demonstrates** parameter $p < 0.08$.

(d) *One-sided tests.*
Right-sided and left-sided tests are (i) **one-sided** (ii) **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 (i) **0.04** (ii) **0.07** (iii) **0.09**.

```
prop1.test(36, 600, 0.08, 0.05, "two.sided") # approx 1-proportion test for p, two-sided

    p.null        p.hat lower z crit upper z crit  z test stat      p value
 0.08000000   0.06000000  -1.95996398   1.95996398  -1.80578780   0.07095149
```

Level of significance $\alpha = $ (i) **0.01** (ii) **0.05** (iii) **0.10**.

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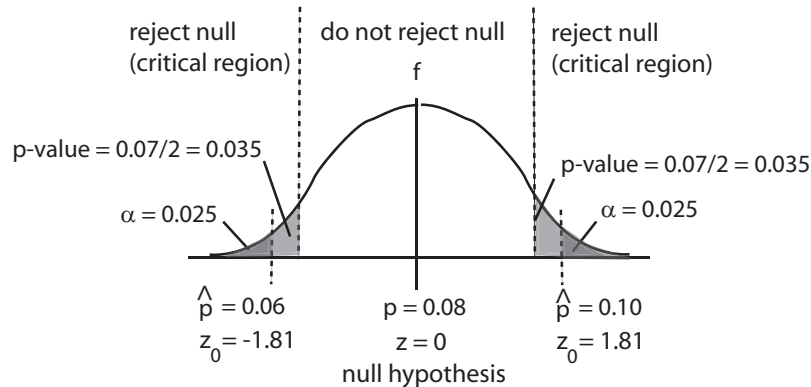

Figure 5.4: Two-sided test

(d) *Claim (final conclusion).* The *claim* is

    i. there is a *change* in defective batteries, so $H_1 : p \neq 0.08$.
    ii. proportion of defective batteries remains the *same*, so $H_0 : p = 0.08$.

    and since the null, $H_0 : p = 0.08$, is *not* rejected,

    i. the data supports this claim, so $H_1 : p \neq 0.08$.
    ii. the data does *not* support this claim, so $H_0 : p = 0.08$.

    The claim (i) **does** (ii) **does not** have an "equals" in it, refers to H_a.

5. *Test p, two-sided, different claim: 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. Assuming the percent defective either *changes* or not, does the
   data support the claim 8% of batteries are defective 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 $H_1 : p \neq 0.08$

(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

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

which equals (i) **0.04**   (ii) **0.07** (iii) **0.09**.

```
prop1.test(36, 600, 0.08, 0.05, "two.sided") # approx 1-proportion test for p, two-sided

     p.null       p.hat lower z crit upper z crit  z test stat      p value
 0.08000000  0.06000000  -1.95996398   1.95996398  -1.80578780   0.07095149
```

Level of significance $\alpha = $ (i) **0.01** (ii) **0.05** (iii) **0.10**.

(c) *Conclusion.*

Since p–value $= 0.07 > \alpha = 0.05$,

(i) **do not reject** (ii) **reject** null guess: $H_0 : p = 0.08$.

So, sample $\hat{p}$ indicates population proportion p

(i) **is less than**   (ii) **equals** (iii) **is greater than** 0.08: $H_0 : p = 0.08$.

(d) *Claim (final conclusion).* The *claim* is

i. there is a *change* in defective batteries, so $H_1 : p \neq 0.08$.

ii. proportion of defective batteries remains the *same*, so $H_0 : p = 0.08$.

and since the null, $H_0 : p = 0.08$, is *not* rejected,

i. is sufficient evidence to reject this claim, so $H_1 : p \neq 0.08$.

ii. is *in*sufficient evidence to reject this claim, so $H_0 : p = 0.08$.

The claim (i) **does**   (ii) **does not** have an "equals" in it, refers to H_0.

(e) *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 (i) **one-sided** (ii) **two-sided** test.

(f) *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 $H_1 : p_{0,2} \leq p \leq p_{0,3}$

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

      i. Statement
      ii. Test
      iii. Conclusions

6. *Power: defection batteries.*
    In a battery factory, $p_0 = 0.08$ of all batteries made are assumed to be defective.
    Calculate and compare power $1 - \beta$ when the actual (true, correct) proportion
    defective is either $p = 0.09$ or $p = 0.12$. Assume $\alpha = 0.05$ when $p_0 = 0.08$.

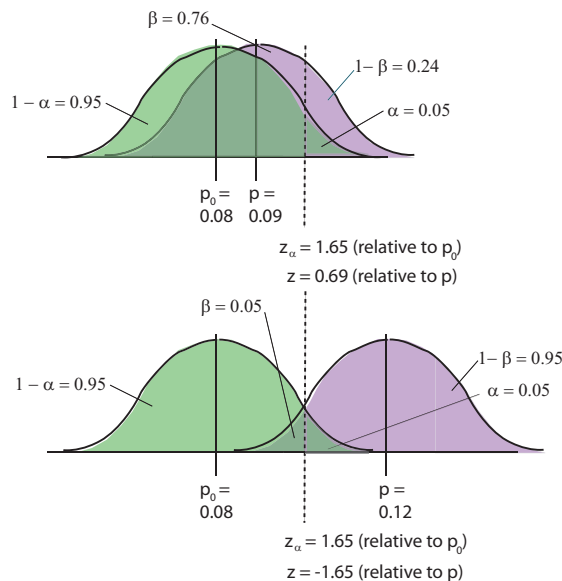| chosen ↓ actual →	null H_0 true	alternative H_1 true
choose null H_0	correct decision	type II error, β
choose alternative H_1	type I error, α	correct decision: power $= 1 - \beta$

Figure 5.5: Errors and correction decisions for two 1-proportion-tests

(a) *Level of significance* $\alpha = P(Z > z_\alpha) = 0.05$ is the chance of
    (i) **mistakenly rejecting $p_0 = 0.08$**
    (ii) **mistakenly rejecting $p = 0.09$**
    (iii) **mistakenly rejecting $p = 0.12$**
    if null $p_0 = 0.08$ is (assumed) true and whether $p = 0.09$ or $p = 0.12$.

(b) *Power* $1 - \beta = P(Z > z) = 0.24$ *when* $p = 0.09$ is the chance of
   (i) **correctly accepting $p_0 = 0.08$**
   (ii) **correctly accepting $p = 0.09$**
   (iii) **correctly accepting $p = 0.12$**
   if alternative $p = 0.09$ is (actually) true and $p_0 = 0.08$ is (actually) wrong.

(c) *Power* $1 - \beta = P(Z > z) = 0.95$ *when* $p = 0.12$ is the chance of
   (i) **correctly accepting $p_0 = 0.08$**
   (ii) **correctly accepting $p = 0.09$**
   (iii) **correctly accepting $p = 0.12$**
   if alternative $p = 0.12$ is (actually) true and $p_0 = 0.08$ is (actually) wrong.

(d) Since power $1 - \beta = P(Z > z) = 0.24$ when $p = 0.09$ is
   (i) **smaller**   (ii) **larger** than
   power $1 - \beta = P(Z > z) = 0.95$ when $p = 0.12$
   there is greater chance of correctly choosing
   (i) **$p = 0.09$**   (ii) **$p = 0.12$**
   when assumed $p_0 = 0.08$ is wrong.

5.2 Testing Claims About a Proportion

Hypothesis test for p from (assumed) binomial has test statistic

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

where we assume a large simple random sample has been chosen, $np_0 \geq 5$ and $n(1 - p_0) \geq 5$. We consider three approaches to testing: p-value, traditional and confidence interval. We also discuss power, $1 - \beta$ and testing for p when sample size is small.

Exercise 5.2 (Testing Claims About a 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 traditional approaches to hypothesis testing.

   (a) *Check assumptions.*
      Since $np_0 = 600(0.08) = 48 > 5$, $n(1 - p_0) = 600(1 - 0.08) = 552 > 5$, assumptions (i) **have**   (ii) **have not** been satisfied, so this large sample test is possible.

   (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

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

(i) **0.00** (ii) **0.09** (iii) **0.18**.

```
prop1.test(54, 600, 0.08, 0.05, "right") # approx 1-proportion test for p

       p.null        p.hat z crit value  z test stat      p value
    0.0800000    0.0900000    1.6448536    0.9028939    0.1832911
```

Level of significance $\alpha =$ (i) **0.01** (ii) **0.05** (iii) **0.10**.

iii. *Conclusion.*

Since p–value $= 0.18 > \alpha = 0.05$,

(i) **do not reject** (ii) **reject** null guess: $H_0 : p = 0.08$.

So, sample $\hat{p}$ indicates population proportion p

(i) **is less than** (ii) **equals** (iii) **is greater than** 0.08: $H_0 : p = 0.08$.

(c) *Traditional 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

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx \frac{0.09 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{600}}} \approx$$

(i) **0.90** (ii) **1.99** (iii) **2.18**.

Critical value $z_\alpha = z_{0.05}$, where $P(Z > z_{0.05}) = 0.05$, is

$$z_\alpha = z_{0.05} \approx$$

(i) **1.28** (ii) **1.65** (iii) **2.58**

```
prop1.test(54, 600, 0.08, 0.05, "right") # approx 1-proportion test for p

       p.null        p.hat z crit value  z test stat      p value
    0.0800000    0.0900000    1.6448536    0.9028939    0.1832911
```

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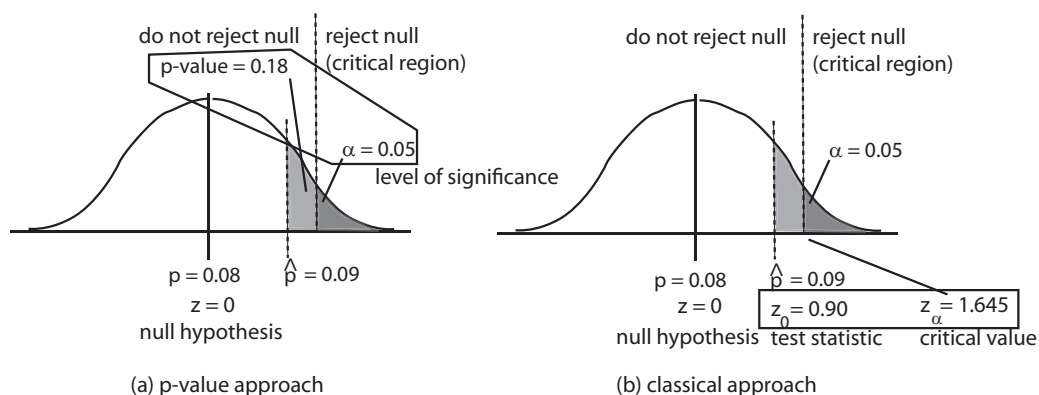

Figure 5.6: P-value versus traditional approach

(d) (i) **True**   (ii) **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 % defective 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 is the *same* as the 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 $np_0 = 600(0.71) = 426 > 5$, $n(1 - p_0) = 600(1 - 0.71) = 174 > 5$, assumptions (i) **have**   (ii) **have not** been satisfied, so this large sample test is possible.

(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

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

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

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

      (i) **0.031** (ii) **0.057** (ii) **0.075**.

      ```
      prop1.test(450, 600, 0.71, 0.05, "two.sided") # approx 1-proportion test for p

                p.null        p.hat lower z crit upper z crit   z test stat        p value
            0.71000000   0.75000000  -1.95996398   1.95996398    2.15927245     0.03082904
      ```

      Level of significance $\alpha = $ **0.01** (ii) **0.05** (iii) **0.10**.

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

   iv. *Claim (final conclusion).* The *claim* is
      A. there is a *change* in proportion overweight, so $H_1 : p \neq 0.71$.
      B. proportion overweight *same* as national proportion, so $H_0 : p = 0.71$.

      and since the null, $H_0 : p = 0.71$, is rejected, there
      A. is sufficient evidence to reject this claim, so $H_1 : p \neq 0.71$.
      B. is *in*sufficient evidence to reject this claim, so $H_0 : p = 0.71$.

The claim (i) **does**   (ii) **does not** have an "equals" in it.

(c) *Traditional 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

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

　　(i) **0.90** (ii) **1.99** (iii) **2.16**.
　　Critical values at $\pm\alpha = 0.05$,

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

　　(i) $\pm$**1.28** (ii) $\pm$**1.65** (ii) $\pm$**1.96**

```
prop1.test(450, 600, 0.71, 0.05, "two.sided") # approx 1-proportion test for p

        p.null         p.hat lower z crit upper z crit  z test stat       p value
    0.71000000    0.75000000  -1.95996398   1.95996398   2.15927245    0.03082904
```

　　iii. *Conclusion.*

　　Since $-z_{0.025} = -1.96 < z_{0.025} = 1.96 < z = 2.16$
　　(test statistic is inside critical region, so "far away" from null 0.71),
　　(i) **do not reject** (ii) **reject** null guess: $H_0 : p = 0.71$.
　　In other words, sample $\hat{p}$ indicates population proportion p
　　(i) **equals** (ii) **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
(i) **(0.018, 0.715)**   (ii) **(0.715, 0.785)**   (iii) **(0.533, 0.867)**.

```
prop1.interval <- function(x,n,conf.level) # function of 1-proportion CI for p
{
  p <- x/n
  z.crit <- -1*qnorm((1-conf.level)/2)
  margin.error <- z.crit*sqrt(p*(1-p)/n)
  ci.lower <- p - margin.error
  ci.upper <- p + margin.error
  dat <- c(p, z.crit, margin.error, ci.lower, ci.upper)
  names(dat) <- c("Mean", "Critical Value", "Margin of Error", "CI lower", "CI upper")
  return(dat)
}
prop1.interval(450,600,0.95) # 1-proportion 95% CI for p

          Mean  Critical Value Margin of Error      CI lower      CI upper
     0.7500000       1.9599640       0.0346476     0.7153524     0.7846476
```

Since null $p_0 = 0.71$ is *out*side CI (0.715, 0.785),
(i) **do not reject** (ii) **reject** null guess: $H_0 : p = 0.71$.
In other words, sample $\hat{p}$ indicates population proportion p
(i) **equals**   (ii) **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 $np_0 = 100(0.065) = 6.5 > 5$, $n(1 - p_0) = 100(1 - 0.065) = 93.5 > 5$, assumptions (i) **have** (ii) **have not** been satisfied (but barely), so compare large-sample z-test with small-sample binomial test.

   (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 (large sample 1-proportion z-test).*
       Chance $\hat{p} = \frac{7}{100} = 0.070$ or more, *if $p_0 = 0.065$,* is

       $$\text{p–value} = P(\hat{p} \geq 0.070) = P\left(\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq \frac{0.070 - 0.065}{\sqrt{\frac{0.065(1-0.065)}{100}}}\right) \approx P\left(Z \geq 0.203\right) \approx$$

       (i) **0.22**   (ii) **0.42**   (iii) **0.48**.

       ```
       prop1.test(7, 100, 0.065, 0.05, "right") # approx 1-proportion 95% test for p

              p.null          p.hat z crit Value  z test stat      p value
           0.0650000      0.0700000    1.6448536    0.2028185    0.4196385
       ```

       Level of significance $\alpha = $ (i) **0.01** (ii) **0.05** (iii) **0.10**.

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

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

       (i) **0.22**   (ii) **0.42**   (iii) **0.48**.

       ```
       1-pbinom(6,100,0.065) # direct calculation of binomial p-value

       [1] 0.4761099
       ```

       Level of significance $\alpha = $ (i) **0.01** (ii) **0.05** (iii) **0.10**.

   (e) *Conclusion.*
       Since p–value = 0.48(or 0.42) $> \alpha = 0.05$,
       (i) **do not reject** (ii) **reject** null guess: $H_0 : p = 0.065$.
       So, sample $\hat{p}$ indicates population proportion p
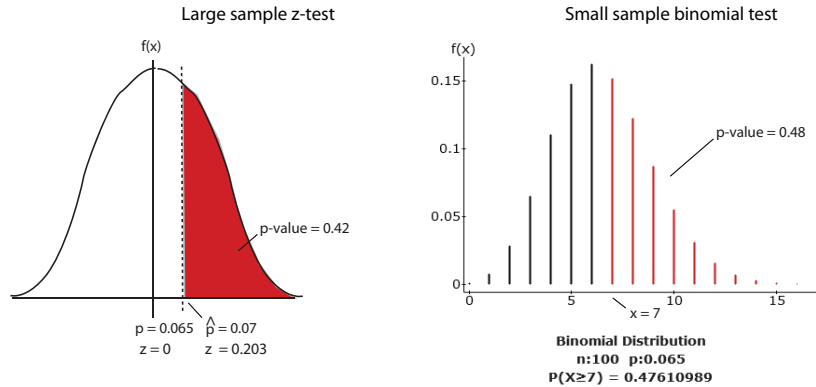   **is less than** (ii) **equals** (iii) **is greater than** 0.065: $H_0 : p = 0.065$.

Figure 5.7: Large sample z-test and small sample binomial test

(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 (i) **null** (ii) **alternative** hypothesis.

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

    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.

4. *Power $1 - \beta$ of the 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. (i) Calculate the *power $1 - \beta$* of the test to detect this, that $p = 0.09$ when null is $p_0 = 0.08$ at $\alpha = 0.05$. (ii) Repeat, but calculate the power of the test to detect $p = 0.12$.

(a) *Power $1 - \beta$ to detect $p = 0.09$ when $p_0 = 0.08$.*
Since, at $\alpha = 0.05$, where $z_\alpha = z_{0.05} = 1.65$

$$z_{0.05} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \Rightarrow \hat{p} = z_{0.05}\sqrt{\frac{p_0(1-p_0)}{n}} + p_0 \approx 1.65\sqrt{\frac{0.08(1-0.08)}{600}} + 0.08 \approx$$

(i) **0.24**   (ii) **0.098**   (iii) **0.95**, the value of the *critical sample proportion*,

and so the power is given by

$$1 - \beta \approx P\left(\hat{p} \geq 0.098\right) \approx P\left(Z \geq \frac{0.098 - 0.09}{\sqrt{\frac{0.09(1-0.09)}{600}}}\right) \approx P(Z \geq 0.69) \approx$$

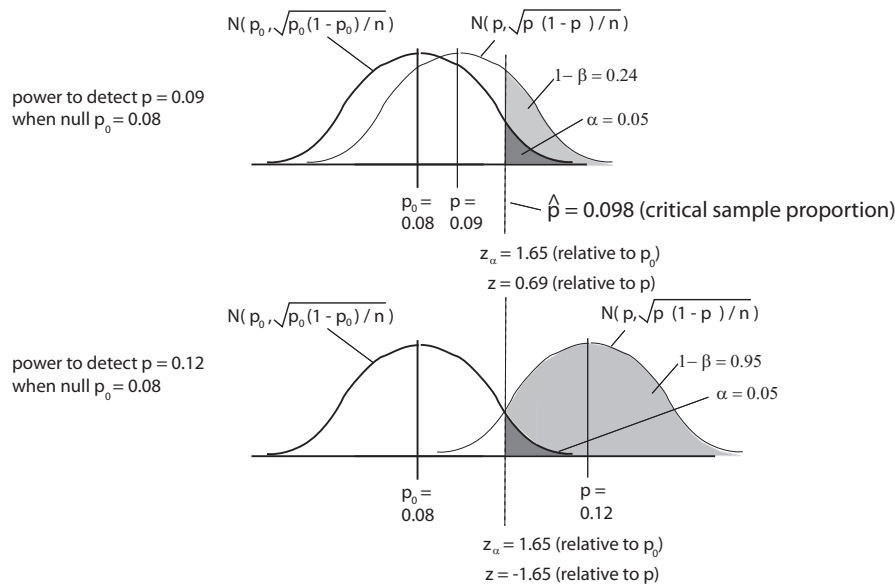(i) **0.24**   (ii) **0.098**   (iii) **0.95**.

Figure 5.8: Power $1 - \beta$ for two 1-proportion-tests

```
prop1.power <- function(n, p.null, p.true, signif.level, type) {
  if(type=="right") {
    z.crit <- -1*qnorm(signif.level)
    p.hat.crit <- z.crit*sqrt(p.null*(1-p.null)/n) + p.null
    power <- 1-pnorm(p.hat.crit,p.true,sqrt(p.true*(1-p.true)/n))
    dat <- c(n, signif.level, p.null, p.true, p.hat.crit, power)
    names(dat) <- c("n", "alpha", "p.null", "p.true", "p.hat.crit.value", "power")
  }
  return(dat)
}
prop1.power(600, 0.08, 0.09, 0.05, "right") # power of test for p
             n           alpha          p.null          p.true p.hat.crit.value            power
   600.00000000      0.05000000      0.08000000      0.09000000       0.09821757       0.24091590
```

(b) *Power $1 - \beta$ to detect $p = 0.12$ when $p_0 = 0.08$.*
Since, at $\alpha = 0.05$, where $z_\alpha = z_{0.05} = 1.65$

$$z_{0.05} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \Rightarrow \hat{p} = z_{0.05}\sqrt{\frac{p_0(1-p_0)}{n}} + p_0 \approx 1.65\sqrt{\frac{0.08(1-0.08)}{600}} + 0.08 \approx$$

(i) **0.24**   (ii) **0.098**   (iii) **0.95**, the value of the *critical sample proportion*,

and so the power is given by

$$1 - \beta \approx P\left(\hat{p} \geq 0.098\right) \approx P\left(Z \geq \frac{0.098 - 0.12}{\sqrt{\frac{0.12(1-0.12)}{600}}}\right) \approx P(Z \geq -1.65) \approx$$

(i) **0.24** (ii) **0.098** (iii) **0.95**

```
prop1.power(600, 0.08, 0.12, 0.05, "right") # power of test for p
```

n	alpha	p.null	p.true	p.hat.crit.value	power
600.00000000	0.05000000	0.08000000	0.12000000	0.09821757	0.94969589

(c) *Comparing powers in two cases.*
Power $1 - \beta$ (0.24) to detect $p = 0.09$ when $p_0 = 0.08$ is
(i) **larger**   (ii) **smaller** than
power $1 - \beta$ (0.95) to detect $p = 0.12$ when $p_0 = 0.08$;
in other words,
the chance to reject $p_0 = 0.08$ when $p = 0.09$ is
(i) **larger**   (ii) **smaller** than
the chance to reject $p_0 = 0.08$ when $p = 0.12$.

5.3   Testing Claims About a Mean

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

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}},$$

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

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

and used when either underlying distribution is normal with no outliers or if random
sample size is large, $n \geq 30$. Typically, more realistic *t*-test used.

Exercise 5.3 (Testing Claims About a Mean)

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

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

which equals (i) **0.18** (ii) **0.20** (iii) **0.23**.

```
mean1.t.test <- function(x.bar,m.null,s,n,signif.level,type) {
  t.test.statistic <- (x.bar-m.null)/(s/sqrt(n))
  if(type=="right") {
    t.crit <- -1*qt(signif.level,n-1)
    p.value <- 1-pt(t.test.statistic,n-1)
  }
  if(type=="left") {
    t.crit <- qt(signif.level,n-1)
    p.value <- pt(t.test.statistic,n-1)
  }
  if(type=="two.sided") {
    t.crit.right <- -1*qt(signif.level/2,n-1)
    t.crit.left <- qt(signif.level/2,n-1)
    t.p.value.right <- 1-pt(t.test.statistic,n-1)
    t.p.value.left <- pt(-1*t.test.statistic,n-1)
    t.p.value <- t.p.value.left + t.p.value.right
  }
  if(type=="two.sided") {
    dat <- c(m.null, x.bar, t.crit.left, t.crit.right, t.test.statistic, t.p.value)
    names(dat) <- c("m.null", "x.bar", "t crit left", "t crit right", "t test stat", "p value")
  } else {
    dat <- c(m.null, x.bar, t.crit, t.test.statistic, p.value)
    names(dat) <- c("m.null", "x.bar", "t crit value", "t test stat", "p value")
  }
  return(dat)
}
mean1.t.test(10.83, 10.05, 3.125, 15, 0.05, "right") # m: mean, s: SD, n: sample size, t-test

      m.null        x.bar t crit value  t test stat      p value
   10.0500000   10.8300000    1.7613101    0.9666966    0.1750497
```

Level of significance $\alpha =$ (i) **0.01** (ii) **0.05** (iii) **0.10**.

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

```
x <- c(27, 41, 22, 27, 23, 35, 30, 33, 24, 27, 28, 22, 24)
```

(a) *Check assumptions (since $n = 13 < 30$): normality.*
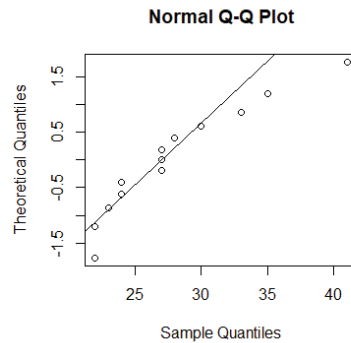
Normal Q-Q Plot

Figure 5.9: Normal probability plot for times

```
qqnorm(x,datax=TRUE) # QQ plot
qqline(x,datax=TRUE) # QQ line
```

Plot indicates activation times (i) **normal** (ii) **not normal** because data veers off the line, but let's continue anyway.

(b) *Statement.*

   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

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

(i) **0.039** (ii) **0.043** (iii) **0.341**.

```
x.bar <- mean(x); m.null <- 25; s <- sqrt(var(x)); n <- length(x)
mean1.t.test(x.bar,m.null,s,n, 0.05, "right") # m: mean, s: SD, n: sample size, t-test
```

```
      m.null         x.bar t crit value  t test stat      p value
 25.00000000   27.92307692    1.78228756    1.87554296   0.04262512
```

Level of significance $\alpha =$ **0.01** (ii) **0.05** (iii) **0.10**.

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

(e) *Related question: unknown σ.*
The σ is **known** (ii) **unknown** in this case
and is approximated by $s \approx$ (i) **5.62** (ii) **5.73** (iii) **5.81**.

<pre style="color:blue">sqrt(var(x))</pre>

<pre>[1] 5.619335</pre>

3. *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

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

      which equals **0.04** (ii)**0.07** (ii) **0.08**.

<pre style="color:blue">mean1.t.test(2.95, 3, 0.18, 30, 0.05, "left") # m: mean, s: SD, n: sample size, t-test</pre>

<pre>   m.null   x.bar t crit value t test stat   p value
   3.00000000   2.95000000 -1.69912703 -1.52145155   0.06948835</pre>

      Level of significance $\alpha =$ (i) **0.01** (ii) **0.05** (iii) **0.10**.

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

(b) *Traditional 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
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{2.95 - 3}{0.18/\sqrt{30}} \approx$$

(i) -1.65 (ii) -1.52 (iii) -1.38.
Critical value at $\alpha = 0.05$,

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

(i) -1.14 (ii) -1.70 (iii) -2.34.

```
mean1.t.test(2.95, 3, 0.18, 30, 0.05, "left") # m: mean, s: SD, n: sample size, t-test

      m.null        x.bar t crit value  t test stat      p value
   3.00000000   2.95000000  -1.69912703  -1.52145155   0.06948835
```

iii. *Conclusion.*

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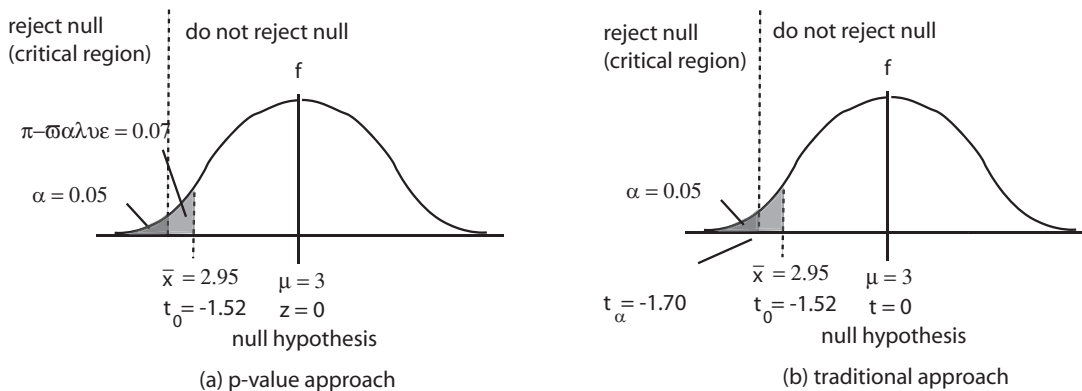

Figure 5.10: Left-sided test of μ: coffee

(c) *Related questions.*

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

    ii. To assume null hypothesis true, means assuming

        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, $\bar{x} = 2.95$.

iii. 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				