

STAT 8010 R Lab 9: Hypothesis Testing

Whitney Huang

October 08, 2020

Contents

Cereal Weight Example	1
Blood Test Example	2
Type I and Type II Errors	3

Cereal Weight Example

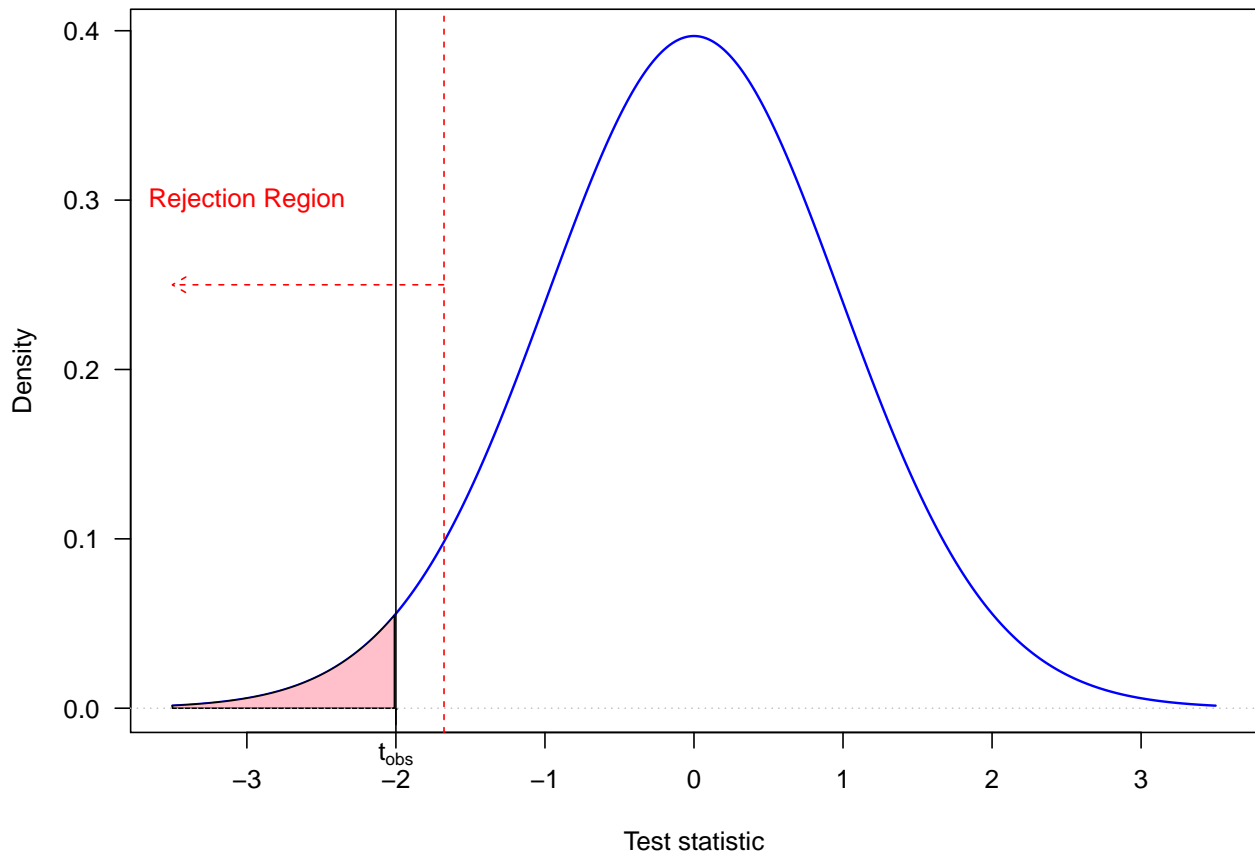
New Age Granola Inc claims that average weight of its cereal boxes is 16 oz. The Genuine Grocery Corporation will send back a shipment if the average weight is any less. Suppose Genuine Grocery Corporation takes a random sample of 49 boxes, weight each one, and compute the sample mean $\bar{X} = 15.9$ oz and sample standard deviation $s = 0.35$ oz. Perform a hypothesis test at 0.05 significant level to determine if they would reject H_0 , and therefore, this shipment

1. $H_0 : \mu = 16$ vs. $H_a : \mu < 16$
2. Test Statistic: $t_{obs} = \frac{15.9-16}{0.35/\sqrt{49}} = -2$
3. Rejection Region Method: $-t_{0.05,48} = -1.68 \Rightarrow$ Rejection Region is $(-\infty, -1.68]$. Since t_{obs} is in rejection region, we reject H_0
4. P-Value Method: $\mathbb{P}(t^* \leq -2) = 0.0256 < \alpha = 0.05 \Rightarrow$ reject H_0
5. Draw a Conclusion: We do have enough statistical evidence to conclude that the average weight is less than 16 oz at 0.05 significant level

```
# P-value
pt(-2, 48)

## [1] 0.02558797

curve(dt(x, 48), from = -3.5, to = 3.5, n = 1001, lwd = 1.5, col = "blue",
      xlab = "Test statistic",
      ylab = "Density", las = 1)
abline(v = -2)
x_grid <- seq(-3.5, 3.5, 0.01)
y_grid <- dt(x_grid, 48)
polygon(c(x_grid[x_grid < -2], rev(x_grid[x_grid < -2])),
        c(y_grid[x_grid < -2], rep(0, length(y_grid[x_grid < -2]))), col = "pink")
abline(h = 0, lty = 3, col = "gray")
axis(1, at = -2, labels = expression(t["obs"]),
     line = -0.75)
abline(v = -qt(0.95, 48), lty = 2, col = "red")
arrows(-1.68, 0.25, x1 = -3.5, length = 0.1,
       lty = 2, col = "red")
text(-3, 0.3, "Rejection Region", col = "red")
```



Blood Test Example

A series of blood tests were run on a particular patient over five days. It is of interest to determine if the mean blood protein for this patient differs from 7.25, the value for healthy adults. Suppose the sample mean ($n = 20$) is 7.35 and sample standard deviation is 0.5. Perform a hypothesis test using significance level of 0.05

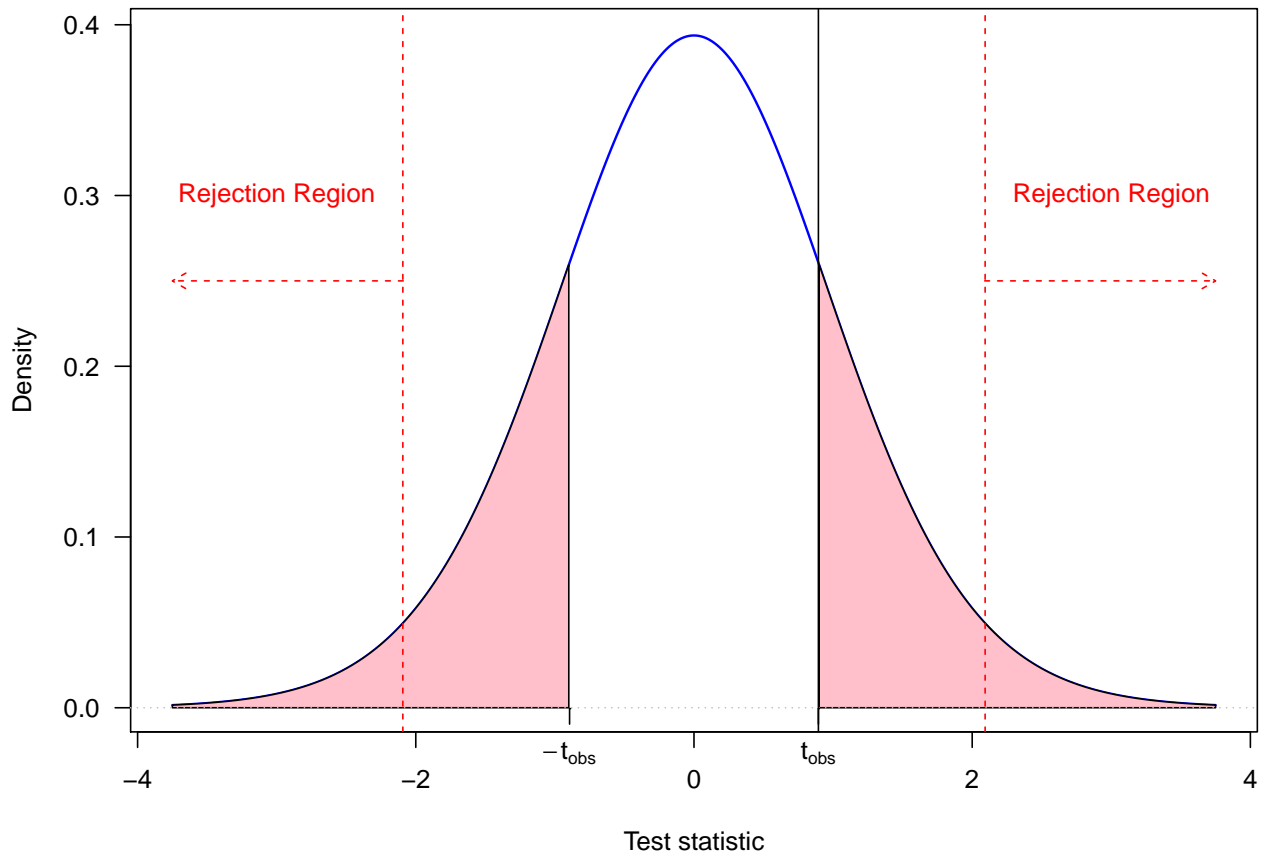
1. $H_0 : \mu = 7.25$ vs. $H_a : \mu \neq 7.25$
2. $t_{obs} = \frac{7.35 - 7.25}{0.5/\sqrt{20}} = 0.8944$
3. P-value: $2 \times \mathbb{P}(t^* \geq 0.8944) = 0.3823 > 0.05$
4. We do not have enough statistical evidence to conclude that the mean blood protein is different from 7.25 at 5% significant level

```
curve(dt(x, 19), from = -3.75, to = 3.75, n = 1001, lwd = 1.5, col = "blue",
      xlab = "Test statistic",
      ylab = "Density", las = 1)
x_grid <- seq(-3.75, 3.75, 0.01)
y_grid <- dt(x_grid, 19)
t_star = 0.8944
polygon(c(x_grid[x_grid < -t_star], rev(x_grid[x_grid < -t_star])),
       c(y_grid[x_grid < -t_star], rep(0, length(y_grid[x_grid < -t_star]))), col = "pink")
polygon(c(x_grid[x_grid > t_star], rev(x_grid[x_grid > t_star])),
       c(y_grid[x_grid > t_star], rep(0, length(y_grid[x_grid > t_star]))), col = "pink")
abline(h = 0, lty = 3, col = "gray")
axis(1, at = 0.8944, labels = expression(t["obs"]), line = -0.75)
```

```

axis(1, at = -0.8944, labels = expression(-t["obs"]), line = -0.75)
abline(v = -qt(0.975, 19), lty = 2, col = "red")
abline(v = qt(0.975, 19), lty = 2, col = "red")
abline(v = 0.8944)
arrows(-2.09, 0.25, x1 = -3.75, length = 0.1,
       lty = 2, col = "red")
arrows(2.09, 0.25, x1 = 3.75, length = 0.1,
       lty = 2, col = "red")
text(-3, 0.3, "Rejection Region", col = "red")
text(3, 0.3, "Rejection Region", col = "red")

```



Type I and Type II Errors

```

library(scales)
curve(dnorm(x), from = -3.75, to = 6.75, n = 1001, lwd = 1.5, col = "gray",
      xlab = "",
      ylab = "", las = 1,
      xaxt = "n", yaxt = "n")
abline(v = 1.96, lty = 2, col = "blue")
x_grid <- seq(-3.75, 6.75, 0.01)
y_grid <- dnorm(x_grid)
polygon(c(x_grid[x_grid > 1.96], rev(x_grid[x_grid > 1.96])),
       c(y_grid[x_grid > 1.96], rep(0, length(y_grid[x_grid > 1.96]))), col = alpha("pink", 0.2))
curve(dnorm(x, mean = 3), from = -3.75, to = 6.75, n = 1001, lwd = 1.5, add = T,
      xlab = "",

```

```

      ylab = "", las = 1)
x_grid <- seq(-3.75, 6.75, 0.01)
y_grid <- dnorm(x_grid, mean = 3)
polygon(c(x_grid[x_grid < 1.96], rev(x_grid[x_grid < 1.96])),
        c(y_grid[x_grid < 1.96], rep(0, length(y_grid[x_grid < 1.96]))), col = alpha("lightblue", 0.2))
abline(v = 0, lty = 2, col = "gray")
abline(v = 3, lty = 2)
axis(1, at = 0, labels = expression(mu[0]))
axis(1, at = 3, labels = expression(mu[a]))
text(2.2, 0.02, expression(alpha), col = "red")
text(1.6, 0.02, expression(beta), col = "blue")

```

