9-10 (p.305)
Type I error, α, is the probability of rejection the null hypothesis when it is actually correct. That means we need to find the probability of getting 2 or more (out of a sample of 10) defectives if the true percentage of defectives, π, is 10%.

$$\begin{aligned}
\alpha = P(X \geq 2) & \\
&= 1 - [P(X=0) + P(X=1)] \\
&= 1 - (0.9)^{10} - (10)(0.1)(0.9)^{9} \\
&\approx 0.2639
\end{aligned}$$

9-12 (p. 305)
(a) Since the test statistic is approximately -3.38 (see previous homework, in which we solved part b), the p value is very small (less than 0.001). So, we have enough evidence to reject the null hypothesis; the difference is statistically discernable.

9-16 (p. 310)
H_0: $\pi = \frac{1}{6}$
H_A: $\pi > \frac{1}{6}$

(a) To reject H_0 with $\alpha = 0.05$, how many aces do you need?

$$\begin{aligned}
\frac{\frac{x}{100} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} &= 1.65 \\
\frac{x}{100} &= \frac{1}{6} + (1.65)\frac{\sqrt{5}}{60} \\
x &\approx 22.82
\end{aligned}$$

To be conservative then, in a sample of 100, you should require 23 aces before rejecting H_0.

(b) The probability of type I error is given by the area to the right of the vertical line (which is drawn at 0.2282) under the leftmost curve (centered at $\frac{1}{6}$). The probability of type II error is given by the area to the left of the vertical line, but under the rightmost curve (centered at 0.25).

(c) Twenty aces isn't enough to really make you favor the alternative hypothesis; it's slightly more likely to happen under H_0 than H_A. I wouldn't be happy with the decision, though, largely because of the friend's experience.

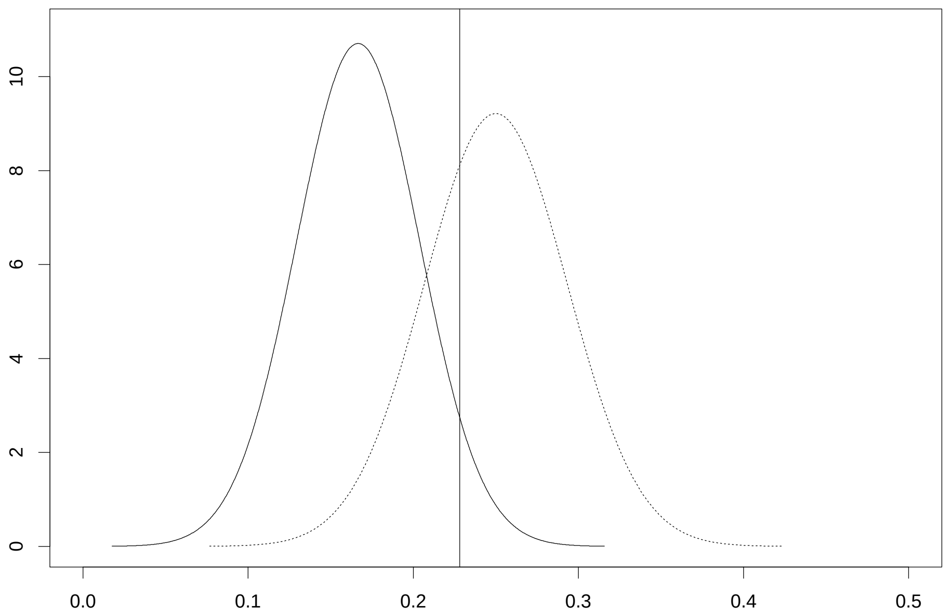

Figure 1: Plot for problem 9-16b

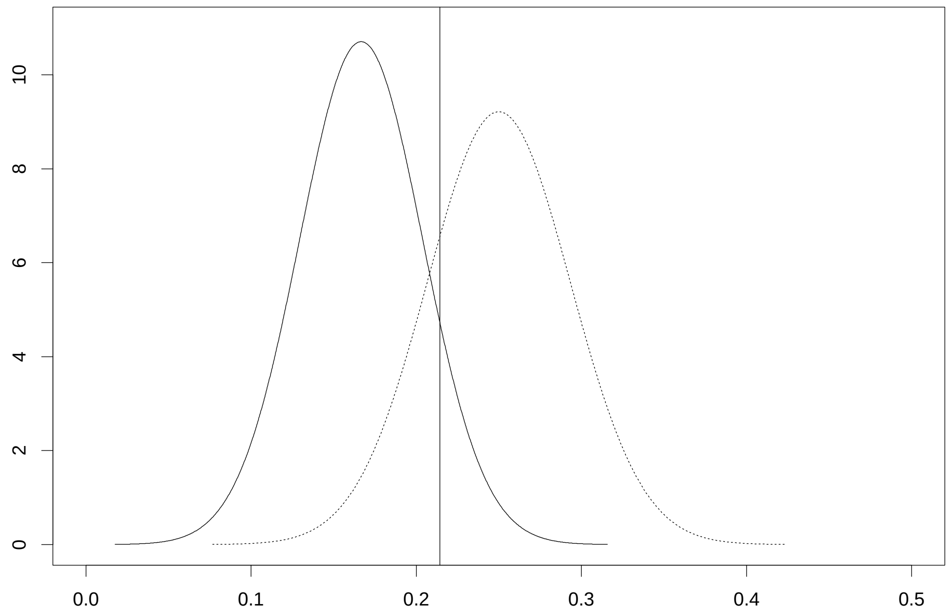

Figure 2: Plot for problem 9-16d

(d) The probability of type I error (about 10%) is given by the area to the right of the vertical line (which is drawn at 0.2144) under the leftmost curve (centered at $\frac{1}{6}$). The probability of type II error (about 0.21) is given by the area to the left of the vertical line, but under the rightmost curve (centered at 0.25).

9-23 (p. 319)
H_0: $\pi_1 - \pi_2 = 0$
H_A: $\pi_1 - \pi_2 \neq 0$

(b) First, we calculate the test statistic:

$$Z = \frac{0.42 - 0.45}{\sqrt{\frac{0.42(0.58)}{1500} + \frac{0.45(0.55)}{1500}}} \approx -1.66$$

Since this is a 2-sided test, $p = 2(0.048) = 0.096$.

(c) (ii) The p value doesn't fall below $\alpha = 0.05$, so we don't have enough evidence to reject the null hypothesis.

9-25 (p. 319)
H_0: $\mu = 14000$
H_A: $\mu > 14000$

(b) First, we calculate the test statistic:

$$Z = \frac{14740 - 14000}{\frac{2000}{\sqrt{25}}} = 1.85$$

Since this is a 1-sided test, $p = 0.032$.

(c) (ii) Since the p value does fall below $\alpha = 0.05$, we have enough evidence to reject the null hypothesis. The new process is discernibly better.

9-30 (p. 321)
H_0: $\pi = 20\%$
H_A: $\pi > 20\%$

In particular, when the process is not in control, the percentage of defectives is $pi = 60\%$.
(a)

$$\begin{aligned}
\alpha &= P(X \geq 3|\pi = 0.20) \\
&= 1 - P(X < 3|\pi = 0.20) \\
&= 1 - (0.8)^{10} - (10)(0.2)(0.8)^9 - \frac{(10)(9)}{2}(0.2)^2(0.8)^8 \\
&\approx 0.3222
\end{aligned}$$

(b)

$$\begin{aligned}
\beta &= P(X < 3|\pi = 0.60) \\
&= (0.4)^{10} + (10)(0.6)(0.4)^9 + \frac{(10)(9)}{2}(0.6)^2(0.4)^8 \\
&\approx 0.0123
\end{aligned}$$

(c) G: process in control
S: sample of 10 has less than 3 defectives (this makes us think the process is in control)

C: total cost of our testing/stopping scheme per day

Event	Cost	P(event)
$G \cap S$	\$100 + 10(\$10)	$0.95(1-\alpha)$
$G \cap S^C$	\$100 + 10(\$10) + \$3000	0.95α
$G^C \cap S$	\$100 + 10(\$10) + \$18000	0.05β
$G^C \cap S^C$	\$100 + 10(\$10)	$0.05(1-\beta)$

We can use the answers in parts a and b above to get the values for α and β. Then, we can calculate the expected daily cost:

$$\begin{aligned} E(C) &\approx 0.6439(\$200) + 0.3061(\$3200) + 0.0006(\$18200) + 0.0494(\$200) \\ &\approx \$1129.10 \end{aligned}$$

(d) If we don't test, we'll have no way of knowing whether the process is "out of whack" or not.

$$E(C) = 0.95(\$0) + 0.05(\$18000) = \$900$$

(e) Must recalculate α and β with the new testing criterion that the process will only be stopped if 4 or more (out of 10) are defective. Then, we can recalculate the expected cost.

$$\begin{aligned} \alpha &= P(X \geq 4 | \pi = 0.20) \\ &= 1 - P(X < 4 | \pi = 0.20) \\ &= 1 - (0.8)^{10} - (10)(0.2)(0.8)^9 \\ &\quad - \frac{(10)(9)}{2}(0.2)^2(0.8)^8 - \frac{(10)(9)(8)}{(3)(2)}(0.2)^3(0.8)^7 \\ &\approx 0.1209 \end{aligned}$$

$$\begin{aligned} \beta &= P(X < 4 | \pi = 0.60) \\ &= (0.4)^{10} + (10)(0.6)(0.4)^9 + \frac{(10)(9)}{2}(0.6)^2(0.4)^8 + \frac{(10)(9)(8)}{(3)(2)}(0.6)^3(0.4)^7 \\ &\approx 0.0548 \end{aligned}$$

$$E(C) \approx \$593$$

We can do the same thing for the strategies of requiring 5 or 6 defectives (out of 10) before stopping the process, resulting in expected costs of \$443 or \$548 (respectively).

So, given these choices, the lowest average daily cost is about \$443. The more defectives (out of 10) we require before closing production, the higher our type II error. The fewer defectives we require, the higher the type I error. However, requiring 5 (out of 10) defectives seems to strike the best balance.

(f) If we start taking samples of size 12, that will change our costs, as well as require us to recalculate various values of α and β. The procedure is the basically the same as before, though.

$$E(C) = 0.95(1-\alpha)(\$220) + 0.95\alpha(\$3220) + 0.05\beta(\$18220) + 0.05(1-\beta)(\$220)$$

If we require 6 (out of 12) defectives before stopping, the expected cost is about \$418. This is the lowest-cost strategy for a sample of this size.

(g) If we require 8 of 18 defective to stop, expected cost is about \$378.

(h) My report would just offer a brief discussion of type I and type II error (which I mentioned above), and conclude by recommending the strategy in part g.

10-4 (p. 335)
H_0: $\mu_1 = \mu_2 = \mu_3$
H_A: at least one of the μs is different than the others

(a) We'll need the overall mean before calculating the rest of the table:

$$\bar{\bar{x}} = \frac{50\bar{x_1} + 50\bar{x_2} + 50\bar{x_3}}{150} = \frac{\bar{x_1} + \bar{x_2} + \bar{x_3}}{3} = 29$$

	Sum Sq.	df	Mean Sq.	F	p
Factor A	18900	2	9450	29.6	<0.001
Residual	47000	147	320		
TOTAL	65900	149			

(b) It is fair to say that given our evidence, the mean incomes are statistically different. However, the F-test doesn't tell us how much which means differ from one another.

10-14 (p. 346)
(a) There are $\binom{8}{2}$ pairs that can be compared.

$$\binom{8}{2} = \frac{(8)(7)}{2} = 28$$

(b) The expected number wrong would be (0.05)(28)=1.4. (Remember the formula $n \times \pi$ from the chapter concerning the binomail theorem?)

(c) $\sqrt{(k-1)F_{.05}^{7,192}}$ replaces $t_{.025}^{192} \approx 1.96$ in each of the confidence intervals.

$$\sqrt{(k-1)F_{.05}^{7,192}} \approx \sqrt{(8-1)(2.17)} \approx 3.90$$

So, the width of the intervals is increased by about

$$2(3.90 s_p \sqrt{\frac{1}{25}+\frac{1}{25}} - 1.96 s_p \sqrt{\frac{1}{25}+\frac{1}{25}}) \approx 1.10 s_p$$