Chapter 10: Hypothesis Testing

10.1 See Definition 10.1.

10.2 Note that Y is binomial with parameters $n = 20$ and p.

a. If the experimenter concludes that less than 80% of insomniacs respond to the drug when actually the drug induces sleep in 80% of insomniacs, a type I error has occurred.

b. $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true}) = P(Y \le 12 \mid p = .8) = .032$ (using Appendix III).

c. If the experimenter does not reject the hypothesis that 80% of insomniacs respond to the drug when actually the drug induces sleep in fewer than 80% of insomniacs, a type II error has occurred.

d. $\beta(.6) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 12 \mid p = .6) = 1 - P(Y \le 12 \mid p = .6) = .416$.

e. $\beta(.4) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 12 \mid p = .4) = .021$.

10.3 **a.** Using the Binomial Table, $P(Y \le 11 \mid p = .8) = .011$, so $c = 11$.
b. $\beta(.6) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 11 \mid p = .6) = 1 - P(Y \le 11 \mid p = .6) = .596$.
c. $\beta(.4) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 11 \mid p = .4) = .057$.

10.4 The parameter p = proportion of ledger sheets with errors.

a. If it is concluded that the proportion of ledger sheets with errors is larger than .05, when actually the proportion is equal to .05, a type I error occurred.

b. By the proposed scheme, H_0 will be rejected under the following scenarios (let E = error, N = no error):

Sheet 1	Sheet 2	Sheet 3
N	*N*	.
N	*E*	*N*
E	*N*	*N*
E	*E*	*N*

With $p = .05$, $\alpha = P(NN) + P(NEN) + P(ENN) + P(EEN) = (.95)^2 + 2(.05)(.95)^2 + (.05)^2(.95) = .995125$.

c. If it is concluded that $p = .05$, but in fact $p > .05$, a type II error occurred.

d. $\beta(p_a) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = \text{P}(EEE, NEE, \text{ or } ENE \mid p_a) = 2p_a^2(1-p_a) + p_a^3$.

10.5 Under H_0, Y_1 and Y_2 are uniform on the interval (0, 1). From Example 6.3, the distribution of $U = Y_1 + Y_2$ is

$$g(u) = \begin{cases} u & 0 \le u \le 1 \\ 2-u & 1 < u \le 2 \end{cases}$$

Test 1: $P(Y_1 > .95) = .05 = \alpha$.

Test 2: $\alpha = .05 = P(U > c) = \int_c^2 (2-u)\,du = 2 = 2c + .5c^2$. Solving the quadratic gives the plausible solution of $c = 1.684$.

10.6 The test statistic Y is binomial with $n = 36$.

a. $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true}) = P(|Y - 18| \geq 4 \mid p = .5) = P(Y \leq 14) + P(Y \geq 22) = .243$.

b. $\beta = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(|Y - 18| \leq 3 \mid p = .7) = P(15 \leq Y \leq 21 \mid p = .7) =$ .09155.

10.7 **a.** False, H_0 is not a statement involving a random quantity.
b. False, for the same reason as part a.
c. True.
d. True.
e. False, this is given by α.
f. **i.** True.
ii. True.
iii. False, β and α behave inversely to each other.

10.8 Let Y_1 and Y_2 have binomial distributions with parameters $n = 15$ and p.

a. $\alpha = P(\text{reject } H_0 \text{ in stage 1} \mid H_0 \text{ true}) + \text{P}(\text{reject } H_0 \text{ in stage 2} \mid H_0 \text{ true})$

$$= P(Y_1 \geq 4) + P(Y_1 + Y_2 \geq 6,\ Y_1 \leq 3) = P(Y_1 \geq 4) + \sum_{i=0}^{3} P(Y_1 + Y_2 \geq 6,\ Y_1 \leq i)$$

$$= P(Y_1 \geq 4) + \sum_{i=0}^{3} P(Y_2 \geq 6 - i)P(Y_1 \leq i) = .0989 \text{ (calculated with } p = .10).$$

Using R, this is found by:

```
> 1 - pbinom(3,15,.1)+sum((1-pbinom(5-0:3,15,.1))*dbinom(0:3,15,.1))
[1] 0.0988643
```

b. Similar to part a with $p = .3$: $\alpha = .9321$.

c. $\beta = P(\text{fail to reject } H_0 \mid p = .3)$

$$= \sum_{i=0}^{3} P(Y_1 = i,\ Y_1 + Y_2 \leq 5) = \sum_{i=0}^{3} P(Y_2 = 5 - i)P(\ Y_1 = i) = .0679.$$

10.9 **a.** The simulation is performed with a known $p = .5$, so rejecting H_0 is a type I error.
b.-e. Answers vary.
f. This is because of part a.
g.-h. Answers vary.

10.10 **a.** An error is the rejection of H_0 (type I).
b. Here, the error is failing to reject H_0 (type II).
c. H_0 is rejected more frequently the further the true value of p is from .5.
d. Similar to part **c**.

10.11 **a.** The error is failing to reject H_0 (type II).
b.-d. Answers vary.

10.12 Since β and α behave inversely to each other, the simulated value for β should be smaller for α = .10 than for α = .05.

10.13 The simulated values of β and α should be closer to the nominal levels specified in the simulation.

10.14 **a.** The smallest value for the test statistic is –.75. Therefore, since the RR is $\{z < -.84\}$, the null hypothesis will never be rejected. The value of n is far too small for this large–sample test.
b. Answers vary.
c. H_0 is rejected when $\hat{p} = 0.00$. $P(Y = 0 \mid p = .1) = .349 > .20$.
d. Answers vary, but n should be large enough.

10.15 **a.** Answers vary.
b. Answers vary.

10.16 **a.** Incorrect decision (type I error).
b. Answers vary.
c. The simulated rejection (error) rate is .000, not close to $\alpha = .05$.

10.17 **a.** H_0: $\mu_1 = \mu_2$, H_a: $\mu_1 > \mu_2$.
b. Reject if $Z > 2.326$, where Z is given in Example 10.7 ($D_0 = 0$).
c. $z = .075$.
d. Fail to reject H_0 – not enough evidence to conclude the mean distance for breaststroke is larger than individual medley.
e. The sample variances used in the test statistic were too large to be able to detect a difference.

10.18 H_0: $\mu = 13.20$, H_a: $\mu < 13.20$. Using the large sample test for a mean, $z = -2.53$, and with $\alpha = .01$, $-z_{.01} = -2.326$. So, H_0 is rejected: there is evidence that the company is paying substandard wages.

10.19 H_0: $\mu = 130$, H_a: $\mu < 130$. Using the large sample test for a mean, $z = \frac{128.6-130}{2.1/\sqrt{40}} = -4.22$ and with $-z_{.05} = -1.645$, H_0 is rejected: there is evidence that the mean output voltage is less than 130.

10.20 H_0: $\mu \geq 64$, H_a: $\mu < 64$. Using the large sample test for a mean, $z = -1.77$, and w/ $\alpha = .01$, $-z_{.01} = -2.326$. So, H_0 is not rejected: there is not enough evidence to conclude the manufacturer's claim is false.

10.21 Using the large–sample test for two means, we obtain $z = 3.65$. With $\alpha = .01$, the test rejects if $|z| > 2.576$. So, we can reject the hypothesis that the soils have equal mean shear strengths.

10.22 **a.** The mean pretest scores should probably be equal, so letting μ_1 and μ_2 denote the mean pretest scores for the two groups, H_0: $\mu_1 = \mu_2$, H_a: $\mu_1 \neq \mu_2$.
b. This is a two–tailed alternative: reject if $|z| > z_{\alpha/2}$.
c. With $\alpha = .01$, $z_{.005} = 2.576$. The computed test statistic is $z = 1.675$, so we fail to reject H_0: we cannot conclude the there is a difference in the pretest mean scores.

10.23 **a.-b.** Let μ_1 and μ_2 denote the mean distances. Since there is no prior knowledge, we will perform the test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, which is a two–tailed test.

c. The computed test statistic is $z = -.954$, which does not lead to a rejection with $\alpha = .10$: there is not enough evidence to conclude the mean distances are different.

10.24 Let p = proportion of overweight children and adolescents. Then, H_0: $p = .15$, H_a: $p < .15$ and the computed large sample test statistic for a proportion is $z = -.56$. This does not lead to a rejection at the $\alpha = .05$ level.

10.25 Let p = proportion of adults who always vote in presidential elections. Then, H_0: $p = .67$, H_a: $p \neq .67$ and the large sample test statistic for a proportion is $|z| = 1.105$. With $z_{.005} = 2.576$, the null hypothesis cannot be rejected: there is not enough evidence to conclude the reported percentage is false.

10.26 Let p = proportion of Americans with brown eyes. Then, H_0: $p = .45$, H_a: $p \neq .45$ and the large sample test statistic for a proportion is $z = -.90$. We fail to reject H_0.

10.27 Define: p_1 = proportion of English–fluent Riverside students
p_2 = proportion of English–fluent Palm Springs students.
To test H_0: $p_1 - p_2 = 0$, versus H_a: $p_1 - p_2 \neq 0$, we can use the large–sample test statistic

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}.$$

However, this depends on the (unknown) values p_1 and p_2. Under H_0, $p_1 = p_2 = p$ (i.e. they are samples from the same binomial distribution), so we can "pool" the samples to estimate p:

$$\hat{p}_p = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{Y_1 + Y_2}{n_1 + n_2}.$$

So, the test statistic becomes

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

Here, the value of the test statistic is $z = -.1202$, so a significant difference cannot be supported.

10.28 **a.** (Similar to 10.27) Using the large–sample test derived in Ex. 10.27, the computed test statistic is $z = -2.254$. Using a two–sided alternative, $z_{.025} = 1.96$ and since $|z| > 1.96$, we can conclude there is a significant difference between the proportions.

b. Advertisers should consider targeting females.

10.29 Note that color A is preferred over B and C if it has the highest probability of being purchased. Thus, let p = probability customer selects color A. To determine if A is preferred, consider the test H_0: $p = 1/3$, H_a: $p > 1/3$. With $\hat{p} = 400/1000 = .4$, the test statistic is $z = 4.472$. This rejects H_0 with $\alpha = .01$, so we can safely conclude that color A is preferred (note that it was assumed that "the first 1000 washers sold" is a random sample).

10.30 Let $\hat{p}$ = sample percentage preferring the product. With $\alpha = .05$, we reject H_0 if

$$\frac{\hat{p} - .2}{\sqrt{.2(.8)/100}} < -1.645 .$$

Solving for $\hat{p}$, the solution is $\hat{p} < .1342$.

10.31 The assumptions are: (1) a random sample (2) a (limiting) normal distribution for the pivotal quantity (3) known population variance (or sample estimate can be used for large n).

10.32 Let p = proportion of U.S. adults who feel the environment quality is fair or poor. To test H_0: $p = .50$ vs. H_a: $p > 50$, we have that $\hat{p} = .54$ so the large–sample test statistic is $z = 2.605$ and with $z_{.05} = 1.645$, we reject H_0 and conclude that there is sufficient evidence to conclude that a majority of the nation's adults think the quality of the environment is fair or poor.

10.33 (Similar to Ex. 10.27) Define:

p_1 = proportion of Republicans strongly in favor of the death penalty
p_2 = proportion of Democrats strongly in favor of the death penalty

To test H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$, we can use the large–sample test derived in Ex. 10.27 with $\hat{p}_1 = .23$, $\hat{p}_2 = .17$, and $\hat{p}_p = .20$. Thus, $z = 1.50$ and for $z_{.05} = 1.645$, we fail to reject H_0: there is not enough evidence to support the researcher's belief.

10.34 Let μ = mean length of stay in hospitals. Then, for H_0: $\mu = 5$, H_a: $\mu > 5$, the large sample test statistic is $z = 2.89$. With $\alpha = .05$, $z_{.05} = 1.645$ so we can reject H_0 and support the agency's hypothesis.

10.35 (Similar to Ex. 10.27) Define:

p_1 = proportion of currently working homeless men
p_2 = proportion of currently working domiciled men

The hypotheses of interest are H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 < 0$, and we can use the large–sample test derived in Ex. 10.27 with $\hat{p}_1 = .30$, $\hat{p}_2 = .38$, and $\hat{p}_p = .355$. Thus, $z = -1.48$ and for $-z_{.01} = -2.326$, we fail to reject H_0: there is not enough evidence to support the claim that the proportion of working homeless men is less than the proportion of working domiciled men.

10.36 (similar to Ex. 10.27) Define:

p_1 = proportion favoring complete protection
p_2 = proportion desiring destruction of nuisance alligators
Using the large–sample test for H_0: $p_1 - p_2 = 0$ versus H_a: $p_1 - p_2 \neq 0$, $z = -4.88$. This value leads to a rejections at the α = .01 level so we conclude that there is a difference.

10.37 With H_0: μ = 130, this is rejected if $z = \frac{\bar{y}-130}{\sigma/\sqrt{n}} < -1.645$, or if $\bar{y} < 130 - \frac{1.645\sigma}{\sqrt{n}} = 129.45$. If μ = 128, then $\beta = P(\bar{Y} > 129.45 \mid \mu = 128) = P(Z > \frac{129.45-128}{2.1/\sqrt{40}}) = P(Z > 4.37) = .0000317$.

10.38 With H_0: μ ≥ 64, this is rejected if $z = \frac{\bar{y}-64}{\sigma/\sqrt{n}} < -2.326$, or if $\bar{y} < 64 - \frac{2.326\sigma}{\sqrt{n}} = 61.36$. If μ = 60, then $\beta = P(\bar{Y} > 61.36 \mid \mu = 60) = P(Z > \frac{61.36-60}{8/\sqrt{50}}) = P(Z > 1.2) = .1151$.

10.39 In Ex. 10.30, we found the rejection region to be: $\{\hat{p} < .1342\}$. For p = .15, the type II error rate is $\beta = P(\hat{p} > .1342 \mid p = .15) = P\left(Z > \frac{.1342-.15}{\sqrt{.15(.85)/100}}\right) = P(Z > -.4424) = .6700$.

10.40 Refer to Ex. 10.33. The null and alternative tests were H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$. We must find a common sample size n such that α = P(reject H_0 | H_0 true) = .05 and β = P(fail to reject H_0 | H_a true) ≤ .20. For α = .05, we use the test statistic

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}}$$ such that we reject H_0 if $Z \geq z_{.05} = 1.645$. In other words,

$$\text{Reject } H_0 \text{ if: } \hat{p}_1 - \hat{p}_2 \geq 1.645\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}.$$

For β, we fix it at the largest acceptable value so $P(\hat{p}_1 - \hat{p}_2 \leq c \mid p_1 - p_2 = .1) = .20$ for some c, or simply

$$\text{Fail to reject } H_0 \text{ if: } \frac{\hat{p}_1 - \hat{p}_2 - .1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}} = -.84, \text{ where } -.84 = z_{.20}.$$

Let $\hat{p}_1 - \hat{p}_2 = 1.645\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}$ and substitute this in the above statement to obtain

$$-.84 = \frac{1.645\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}} - .1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}} = 1.645 - \frac{.1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}}, \text{ or simply } 2.485 = \frac{.1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}}.$$

Using the hint, we set $p_1 = p_2 = .5$ as a "worse case scenario" and find that

$$2.485 = \frac{.1}{\sqrt{.5(.5)\left[\frac{1}{n} + \frac{1}{n}\right]}}.$$

The solution is $n = 308.76$, so the common sample size for the researcher's test should be $n = 309$.

10.41 Refer to Ex. 10.34. The rejection region, written in terms of $\bar{y}$, is

$$\left\{\frac{\bar{y}-5}{3.1/\sqrt{500}} > 1.645\right\} \Leftrightarrow \{\bar{y} > 5.228\}.$$

Then, $\beta = P(\bar{y} \leq 5.228 \mid \mu = 5.5) = P\left(Z \leq \frac{5.228-5.5}{3.1/\sqrt{500}}\right) = P(Z \leq 1.96) = .025$.

10.42 Using the sample size formula given in this section, we have

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2} = 607.37,$$

so a sample size of 608 will provide the desired levels.

10.43 Let μ_1 and μ_2 denote the mean dexterity scores for those students who did and did not (respectively) participate in sports.

a. For H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$ with $\alpha = .05$, the rejection region is $\{z > 1.645\}$ and the computed test statistic is

$$z = \frac{32.19 - 31.68}{\sqrt{\frac{(4.34)^2}{37} + \frac{(4.56)^2}{37}}} = .49.$$

Thus H_0 is not rejected: there is insufficient evidence to indicate the mean dexterity score for students participating in sports is larger.

b. The rejection region, written in terms of the sample means, is

$$\bar{Y}_1 - \bar{Y}_2 > 1.645\sqrt{\frac{(4.34)^2}{37} + \frac{(4.56)^2}{37}} = 1.702.$$

Then, $\beta = P(\bar{Y}_1 - \bar{Y}_2 \le 1.702 \mid \mu_1 - \mu_2 = 3) = P\left(Z \le \frac{1.703-3}{\hat{\sigma}_{\bar{Y}_1 - \bar{Y}_2}}\right) = P(Z < -1.25) = .1056.$

10.44 We require $\alpha = P(\bar{Y}_1 - \bar{Y}_2 > c \mid \mu_1 - \mu_2 = 0) = P\left(Z > \frac{c-0}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}\right)$, so that $z_\alpha = \frac{c\sqrt{n}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$. Also,

$\beta = P(\bar{Y}_1 - \bar{Y}_2 \le c \mid \mu_1 - \mu_2 = 3) = P\left(Z \le \frac{(c-3)\sqrt{n}}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$, so that $-z_\beta = \frac{(c-3)\sqrt{n}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$. By eliminating c

in these two expressions, we have $z_\alpha\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}} = 3 - z_\beta\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}$. Solving for n, we have

$$n = \frac{2(1.645)^2[(4.34)^2 + (4.56)^2]}{3^2} = 47.66.$$

A sample size of 48 will provide the required levels of α and β.

10.45 The 99% CI is $1.65 - 1.43 \pm 2.576\sqrt{\frac{(.26)^2}{30} + \frac{(.22)^2}{35}} = .22 \pm .155$ or (.065, .375). Since the interval does not contain 0, the null hypothesis should be rejected (same conclusion as Ex. 10.21).

10.46 The rejection region is $\frac{\hat{\theta} - \theta_0}{\hat{\sigma}_\theta} > z_\alpha$, which is equivalent to $\theta_0 < \hat{\theta} - z_\alpha \hat{\sigma}_{\hat{\theta}}$. The left–hand side is the 100(1 – α)% lower confidence bound for θ.

10.47 (Refer to Ex. 10.32) The 95% lower confidence bound is $.54 - 1.645\sqrt{\frac{.54(.46)}{1060}} = .5148$. Since the value $p = .50$ is less than this lower bound, it does not represent a plausible value for p. This is equivalent to stating that the hypothesis H_0: $p = .50$ should be rejected.

10.48 (Similar to Ex. 10.46) The rejection region is $\frac{\hat{\theta}-\theta_0}{\hat{\sigma}_\theta} < -z_\alpha$, which is equivalent to $\theta_0 > \hat{\theta} + z_\alpha \hat{\sigma}_{\hat{\theta}}$. The left–hand side is the 100(1 – α)% upper confidence bound for θ.

10.49 (Refer to Ex. 10.19) The upper bound is $128.6 + 1.645\left(\frac{2.1}{\sqrt{40}}\right) = 129.146$. Since this bound is less than the hypothesized value of 130, H_0 should be rejected as in Ex. 10.19.

10.50 Let μ = mean occupancy rate. To test H_0: μ ≥ .6, H_a: μ < .6, the computed test statistic is $z = \frac{.58-.6}{.11/\sqrt{120}} = -1.99$. The p–value is given by $P(Z < -1.99) = .0233$. Since this is less than the significance level of .10, H_0 is rejected.

10.51 To test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 represent the two mean reading test scores for the two methods, the computed test statistic is

$$z = \frac{74-71}{\sqrt{\frac{9^2}{50}+\frac{10^2}{50}}} = 1.58.$$

The p–value is given by $P(|Z| > 1.58) = 2P(Z > 1.58) = .1142$, and since this is larger than α = .05, we fail to reject H_0.

10.52 The null and alternative hypotheses are H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$, where p_1 and p_2 correspond to normal cell rates for cells treated with .6 and .7 (respectively) concentrations of actinomycin D.

a. Using the sample proportions .786 and .329, the test statistic is (refer to Ex. 10.27) $z = \frac{.786-.329}{\sqrt{(.557)(.443)\frac{2}{70}}} = 5.443$. The p–value is $P(Z > 5.443) \approx 0$.

b. Since the p–value is less than .05, we can reject H_0 and conclude that the normal cell rate is lower for cells exposed to the higher actinomycin D concentration.

10.53 **a.** The hypothesis of interest is H_0: $\mu_1 = 3.8$, H_a: $\mu_1 < 3.8$, where μ_1 represents the mean drop in FVC for men on the physical fitness program. With $z = -.996$, we have p–value $= P(Z < -1) = .1587$.

b. With α = .05, H_0 cannot be rejected.

c. Similarly, we have H_0: $\mu_2 = 3.1$, H_a: $\mu_2 < 3.1$. The computed test statistic is $z = -1.826$ so that the p–value is $P(Z < -1.83) = .0336$.

d. Since α = .05 is greater than the p–value, we can reject the null hypothesis and conclude that the mean drop in FVC for women is less than 3.1.

10.54 **a.** The hypotheses are H_0: $p = .85$, H_a: $p > .85$, where p = proportion of right–handed executives of large corporations. The computed test statistic is $z = 5.34$, and with α = .01,

$z_{.01} = 2.326$. So, we reject H_0 and conclude that the proportion of right–handed executives at large corporations is greater than 85%

b. Since p–value = $P(Z > 5.34) < .000001$, we can safely reject H_0 for any significance level of .000001 or more. This represents strong evidence against H_0.

10.55 To test H_0: $p = .05$, H_a: $p < .05$, with $\hat{p} = 45/1124 = .040$, the computed test statistic is $z = -1.538$. Thus, p–value = $P(Z < -1.538) = .0616$ and we fail to reject H_0 with $\alpha = .01$. There is not enough evidence to conclude that the proportion of bad checks has decreased from 5%.

10.56 To test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$, where μ_1, μ_2 represent the two mean recovery times for treatments {no supplement} and {500 mg Vitamin C}, respectively. The computed test statistic is $z = \frac{6.9-5.8}{\sqrt{[(2.9)^2+(1.2)^2]/35}} = 2.074$. Thus, p–value = $P(Z > 2.074)$ = .0192 and so the company can reject the null hypothesis at the .05 significance level conclude the Vitamin C reduces the mean recovery times.

10.57 Let p = proportion who renew. Then, the hypotheses are H_0: $p = .60$, H_a: $p \neq .60$. The sample proportion is $\hat{p} = 108/200 = .54$, and so the computed test statistic is $z = -1.732$. The p–value is given by $2P(Z < -1.732) = .0836$.

10.58 The null and alternative hypotheses are H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$, where p_1 and p_2 correspond to, respectively, the proportions associated with groups A and B. Using the test statistic from Ex. 10.27, its computed value is $z = \frac{.74-.46}{\sqrt{.6(.4)\frac{2}{50}}} = 2.858$. Thus, p–value = $P(Z > 2.858) = .0021$. With $\alpha = .05$, we reject H0 and conclude that a greater fraction feel that a female model used in an ad increases the perceived cost of the automobile.

10.59 **a.-d.** Answers vary.

10.60 **a.-d.** Answers vary.

10.61 If the sample size is small, the test is only appropriate if the random sample was selected from a normal population. Furthermore, if the population is not normal and σ is unknown, the estimate s should only be used when the sample size is large.

10.62 For the test statistic to follow a t–distribution, the random sample should be drawn from a normal population. However, the test does work satisfactorily for similar populations that possess mound–shaped distributions.

10.63 The sample statistics are $\bar{y} = 795$, $s = 8.337$.

a. The hypotheses to be tested are H_0: $\mu = 800$, H_a: $\mu < 800$, and the computed test statistic is $t = \frac{795-800}{8.337/\sqrt{5}} = -1.341$. With $5 - 1 = 4$ degrees of freedom, $-t_{.05} = -2.132$ so we fail to reject H_0 and conclude that there is not enough evidence to conclude that the process has a lower mean yield.

b. From Table 5, we find that p–value > .10 since $-t_{.10} = -1.533$.
c. Using the Applet, p–value = .1255.
d. The conclusion is the same.

10.64 The hypotheses to be tested are H_0: $\mu = 7$, H_a: $\mu \neq 7$, where μ = mean beverage volume.

a. The computed test statistic is $t = \frac{7.1-7}{.12/\sqrt{10}} = 2.64$ and with 10 –1 = 9 degrees of freedom, we find that $t_{.025} = 2.262$. So the null hypothesis could be rejected if $\alpha = .05$ (recall that this is a two–tailed test).
b. Using the Applet, $2P(T > 2.64) = 2(.01346) = .02692$.
c. Reject H_0.

10.65 The sample statistics are $\bar{y} = 39.556$, $s = 7.138$.

a. To test H_0: $\mu = 45$, H_a: $\mu < 45$, where μ = mean cost, the computed test statistic is $t = -3.24$. With 18 – 1 = 17 degrees of freedom, we find that $-t_{.005} = -2.898$, so the p–value must be less than .005.
b. Using the Applet, $P(T < -3.24) = .00241$.
c. Since $t_{.025} = 2.110$, the 95% CI is $39.556 \pm 2.11\left(\frac{7.138}{\sqrt{18}}\right)$ or (36.006, 43.106).

10.66 The sample statistics are $\bar{y} = 89.855$, s = 14.904.

a. To test H_0: $\mu = 100$, H_a: $\mu < 100$, where μ = mean DL reading for current smokers, the computed test statistic is $t = -3.05$. With 20 – 1 = 19 degrees of freedom, we find that $-t_{.01} = -2.539$, so we reject H_0 and conclude that the mean DL reading is less than 100.
b. Using Appendix 5, $-t_{.005} = -2.861$, so p–value < .005.
c. Using the Applet, $P(T < -3.05) = .00329$.

10.67 Let μ = mean calorie content. Then, we require H_0: $\mu = 280$, H_a: $\mu > 280$.

a. The computed test statistic is $t = \frac{358-280}{54/\sqrt{10}} = 4.568$. With 10 – 1 = 9 degrees of freedom, $t_{.01} = 2.821$ so H_0 can be rejected: it is apparent that the mean calorie content is greater than advertised.
b. The 99% lower confidence bound is $358 - 2.821\frac{54}{\sqrt{10}} = 309.83$ cal.
c. Since the value 280 is below the lower confidence bound, it is unlikely that $\mu = 280$ (same conclusion).

10.68 The random samples are drawn independently from two normal populations with common variance.

10.69 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$.

a. The computed test statistic is, where $s_p^2 = \frac{10(52)+13(71)}{23} = 62.74$, is given by

$$t = \frac{64-69}{\sqrt{62.74\left(\frac{1}{11}+\frac{1}{14}\right)}} = -1.57.$$

i. With 11 + 14 – 2 = 23 degrees of freedom, $-t_{.10} = -1.319$ and $-t_{.05} = -1.714$. Thus, since we have a two–sided alternative, .10 < p–value < .20.

ii. Using the Applet, $2P(T < -1.57) = 2(.06504) = .13008$.

b. We assumed that the two samples were selected independently from normal populations with common variance.

c. Fail to reject H_0.

10.70 **a.** The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$. The computed test statistic is t = 2.97 (here, $s_p^2 = .0001444$). With 21 degrees of freedom, $t_{.05} = 1.721$ so we reject H_0.

b. For this problem, the hypotheses are H_0: $\mu_1 - \mu_2 = .01$ vs. H_a: $\mu_1 - \mu_2 > .01$. Then, $t = \frac{(.041-.026)-.01}{\sqrt{s_p^2\left(\frac{1}{9}+\frac{1}{12}\right)}} = .989$ and p–value > .10. Using the Applet, $P(T > .989) = .16696$.

10.71 **a.** The summary statistics are: $\bar{y}_1 = 97.856$, $s_1^2 = .3403$, $\bar{y}_2 = 98.489$, $s_2^2 = .3011$. To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = -2.3724$ with 16 degrees of freedom. We have that $-t_{.01} = -2.583$, $-t_{.025} = -2.12$, so $.02 < p\text{–value} < .05$.

b. Using the Applet, $2P(T < -2.3724) = 2(.01527) = .03054$.

R output:
```
> t.test(temp~sex,var.equal=T)

        Two Sample t-test

data:  temp by sex
t = -2.3724, df = 16, p-value = 0.03055
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.19925448 -0.06741219
sample estimates:
mean in group 1 mean in group 2
       97.85556        98.48889
```

10.72 To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = 1.655$ with 38 degrees of freedom. Since we have that $\alpha = .05$, $t_{.025} \approx z_{.025} = 1.96$ so fail to reject H_0 and p–value = $2P(T > 1.655)$ = $2(.05308) = .10616$.

10.73 **a.** To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = 1.92$ with 18 degrees of freedom. Since we have that $\alpha = .05$, $t_{.025} = 2.101$ so fail to reject H_0 and p–value = $2P(T > 1.92)$ = $2(.03542) = .07084$.

b. To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = .365$ with 18 degrees of freedom. Since we have that $\alpha = .05$, $t_{.025} = 2.101$ so fail to reject H_0 and p–value = $2P(T > .365)$ = $2(.35968) = .71936$.

10.74 The hypotheses are H_0: $\mu = 6$ vs. H_a: $\mu < 6$ and the computed test statistic is $t = 1.62$ with 11 degrees of freedom (note that here $\bar{y} = 9$, so H_0 could never be rejected). With α = .05, the critical value is $-t_{.05} = -1.796$ so fail to reject H_0.

10.75 Define μ = mean trap weight. The sample statistics are $\bar{y}$ = 28.935, s = 9.507. To test H_0: $\mu = 30.31$ vs. H_a: $\mu < 30.31$, $t = -.647$ with 19 degrees of freedom. With $\alpha = .05$, the critical value is $-t_{.05} = -1.729$ so fail to reject H_0: we cannot conclude that the mean trap weight has decreased. R output:

```
> t.test(lobster,mu=30.31, alt="less")

        One Sample t-test

data:  lobster
t = -0.6468, df = 19, p-value = 0.2628
alternative hypothesis: true mean is less than 30.31
95 percent confidence interval:
     -Inf 32.61098
```

10.76 **a.** To test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$, where μ_1, μ_2 represent mean plaque measurements for the control and antiplaque groups, respectively.

b. The pooled sample variance is $s_p^2 = \frac{6(.32)^2+6(.32)^2}{12} = .1024$ and the computed test statistic is $t = \frac{1.26-.78}{\sqrt{.1024\left(\frac{2}{7}\right)}} = 2.806$ with 12 degrees of freedom. Since $\alpha = .05$, $t_{.05} = 1.782$ and H_0 is rejected: there is evidence that the antiplaque rinse reduces the mean plaque measurement.

c. With $t_{.01} = 2.681$ and $t_{.005} = 3.005$, $.005 < p$–value $< .01$ (exact: .00793).

10.77 **a.** To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean verbal SAT scores for students intending to major in engineering and language (respectively), the pooled sample variance is $s_p^2 = \frac{14(42)^2+14(45)^2}{28} = 1894.5$ and the computed test statistic is $t = \frac{446-534}{\sqrt{1894.5\left(\frac{2}{15}\right)}} = -5.54$ with 28 degrees of freedom. Since $-t_{.005} = -2.763$, we can reject H_0 and p–value $< .01$ (exact: 6.35375e-06).

b. Yes, the CI approach agrees.

c. To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean math SAT scores for students intending to major in engineering and language (respectively), the pooled sample variance is $s_p^2 = \frac{14(57)^2+14(52)^2}{28} = 2976.5$ and the computed test statistic is $t = \frac{548-517}{\sqrt{2976.5\left(\frac{2}{15}\right)}} = 1.56$ with 28 degrees of freedom. From Table 5, $.10 < p$–value $< .20$ (exact: 0.1299926).

d. Yes, the CI approach agrees.

10.78 **a.** We can find $P(Y > 1000) = P(Z > \frac{1000-800}{40}) = P(Z > 5) \approx 0$, so it is very unlikely that the force is greater than 1000 lbs.

b. Since $n = 40$, the large–sample test for a mean can be used: H_0: $\mu = 800$ vs. H_a: $\mu > 800$ and the test statistic is $z = \frac{825-800}{\sqrt{2350/40}} = 3.262$. With p–value = $P(Z > 3.262) < .00135$, we reject H_0.

c. Note that if $\sigma = 40$, $\sigma^2 = 1600$. To test: H_0: $\sigma^2 = 1600$ vs. H_a: $\sigma^2 > 1600$. The test statistic is $\chi^2 = \frac{39(2350)}{1600} = 57.281$. With $40 - 1 = 39$ degrees of freedom (approximated with 40 degrees of freedom in Table 6), $\chi^2_{.05} = 55.7585$. So, we can reject H_0 and conclude there is sufficient evidence that σ exceeds 40.

10.79 **a.** The hypotheses are: H_0: $\sigma^2 = .01$ vs. H_a: $\sigma^2 > .01$. The test statistic is $\chi^2 = \frac{7(.018)}{.01} =$ 12.6 with 7 degrees of freedom. With $\alpha = .05$, $\chi^2_{.05} = 14.07$ so we fail to reject H_0. We must assume the random sample of carton weights were drawn from a normal population.

b. i. Using Table 6, $.05 < p$–value $< .10$.
ii. Using the Applet, $P(\chi^2 > 12.6) = .08248$.

10.80 The two random samples must be independently drawn from normal populations.

10.81 For this exercise, refer to Ex. 8.125.

a. The rejection region is $\left\{S_1^2 \big/ S_2^2 > F^{\nu_1}_{\nu_2,\alpha/2}\right\} \cup \left\{S_1^2 \big/ S_2^2 < \left(F^{\nu_2}_{\nu_1,\alpha/2}\right)^{-1}\right\}$. If the reciprocal is taken in the second inequality, we have $S_2^2 \big/ S_1^2 > F^{\nu_1}_{\nu_2,\alpha/2}$.

b. $P\left(S_L^2 \big/ S_S^2 > F^{\nu_L}_{\nu_S,\alpha/2}\right) = P\left(S_1^2 \big/ S_2^2 > F^{\nu_1}_{\nu_2,\alpha/2}\right) + P\left(S_2^2 \big/ S_1^2 > F^{\nu_2}_{\nu_1,\alpha/2}\right) = \alpha$, by part a.

10.82 **a.** Let σ_1^2, σ_2^2 denote the variances for compartment pressure for resting runners and cyclists, respectively. To test H_0: $\sigma_1^2 = \sigma_2^2$ vs. H_a: $\sigma_1^2 \neq \sigma_2^2$, the computed test statistic is $F = (3.98)^2/(3.92)^2 = 1.03$. With $\alpha = .05$, $F^9_{9,.025} = 4.03$ and we fail to reject H_0.

b. i. From Table 7, p–value $> .1$.
ii. Using the Applet, $2P(F > 1.03) = 2(.4828) = .9656$.

c. Let σ_1^2, σ_2^2 denote the population variances for compartment pressure for 80% maximal O_2 consumption for runners and cyclists, respectively. To test H_0: $\sigma_1^2 = \sigma_2^2$ vs. H_a: $\sigma_1^2 \neq \sigma_2^2$, the computed test statistic is $F = (16.9)^2/(4.67)^2 = 13.096$ and we reject H_0: the is sufficient evidence to claim a difference in variability.

d. i. From Table 7, p–value $< .005$.
ii. Using the Applet, $2P(F > 13.096) = 2(.00036) = .00072$.

10.83 **a.** The manager of the dairy is concerned with determining if there is a *difference* in the two variances, so a two–sided alternative should be used.

b. The salesman for company *A* would prefer $H_a: \sigma_1^2 < \sigma_2^2$, since if this hypothesis is accepted, the manager would choose company *A*'s machine (since it has a smaller variance).

c. For similar logic used in part b, the salesman for company B would prefer $H_a: \sigma_1^2 > \sigma_2^2$.

10.84 Let σ_1^2, σ_2^2 denote the variances for measurements corresponding to 95% ethanol and 20% bleach, respectively. The desired hypothesis test is $H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_a: \sigma_1^2 \neq \sigma_2^2$ and the computed test statistic is $F = \left(2.78095/.17143\right) = 16.222$.

a. i. With 14 numerator and 14 denominator degrees of freedom, we can approximate the critical value in Table 7 by $F_{14,.005}^{15} = 4.25$, so *p*–value < .01 (two–tailed test).

ii. Using the Applet, $2P(F > 16.222) \approx 0$.

b. We would reject H_0 and conclude the variances are different.

10.85 Since $(.7)^2 = .49$, the hypotheses are: H_0: $\sigma^2 = .49$ vs. H_a: $\sigma^2 > .49$. The sample variance $s^2 = 3.667$ so the computed test statistic is $\chi^2 = \frac{3(3.667)}{.49} = 22.45$ with 3 degrees of freedom. Since $\chi_{.05}^2 = 12.831$, *p*–value < .005 (exact: .00010).

10.86 The hypotheses are: H_0: $\sigma^2 = 100$ vs. H_a: $\sigma^2 > 100$. The computed test statistic is $\chi^2 = \frac{19(144)}{100} = 27.36$. With $\alpha = .01$, $\chi_{.01}^2 = 36.1908$ so we fail to reject H_0: there is not enough evidence to conclude the variability for the new test is higher than the standard.

10.87 Refer to Ex. 10.87. Here, the test statistic is $(.017)^2/(.006)^2 = 8.03$ and the critical value is $F_{12,.05}^{9} = 2.80$. Thus, we can support the claim that the variance in measurements of DDT levels for juveniles is greater than it is for nestlings.

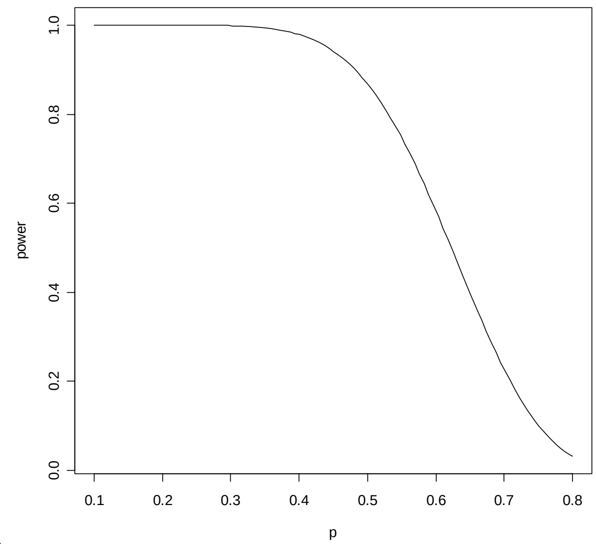
10.88 Refer to Ex. 10.2. Table 1 in Appendix III is used to find the binomial probabilities.

a. power(.4) = $P(Y \leq 12 \mid p = .4) = .979$.

b. power(.5) = $P(Y \leq 12 \mid p = .5) = .86$

c. power(.6) = $P(Y \leq 12 \mid p = .6) = .584$.

d. power(.7) = $P(Y \leq 12 \mid p = .7) = .228$



e. The power function is above.

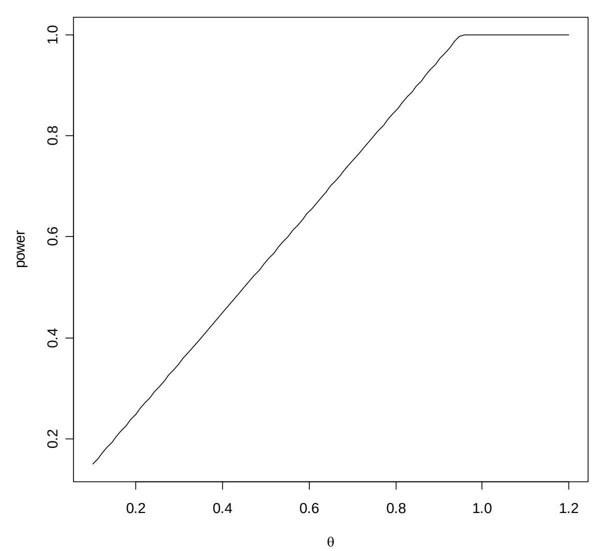
10.89 Refer to Ex. 10.5: $Y_1 \sim \text{Unif}(\theta, \theta + 1)$.

a. $\theta = .1$, so $Y_1 \sim \text{Unif}(.1, 1.1)$ and power(.1) = $P(Y_1 > .95) = \int_{.95}^{1.1} dy = .15$

b. $\theta = .4$: power(.4) = $P(Y > .95) = .45$

c. $\theta = .7$: power(.7) = $P(Y > .95) = .75$

d. $\theta = 1$: power(1) = $P(Y > .95) = 1$



e. The power function is above.

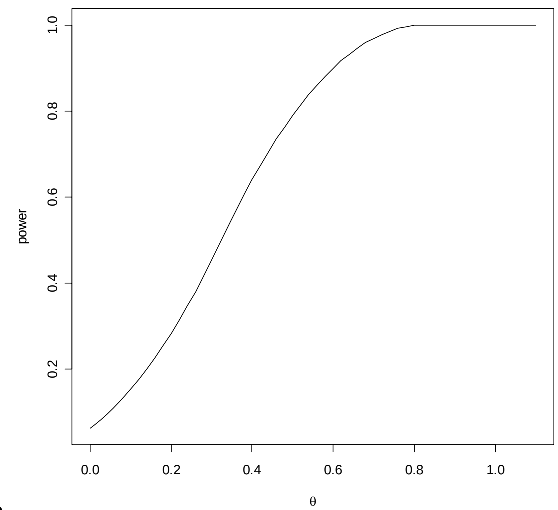
10.90 Following Ex. 10.5, the distribution function for Test 2, where $U = Y_1 + Y_2$, is

$$F_U(u) = \begin{cases} 0 & u < 0 \\ .5u^2 & 0 \le u \le 1 \\ 2u - .5u^2 - 1 & 1 < u \le 2 \\ 1 & u > 2 \end{cases}.$$

The test rejects when $U > 1.684$. The power function is given by:

$$\text{power}(\theta) = P_\theta(Y_1 + Y_2 > 1.684) = P(Y_1 + Y_2 - 2\theta > 1.684 - 2\theta)$$
$$= P(U > 1.684 - 2\theta) = 1 - F_U(1.684 - 2\theta).$$

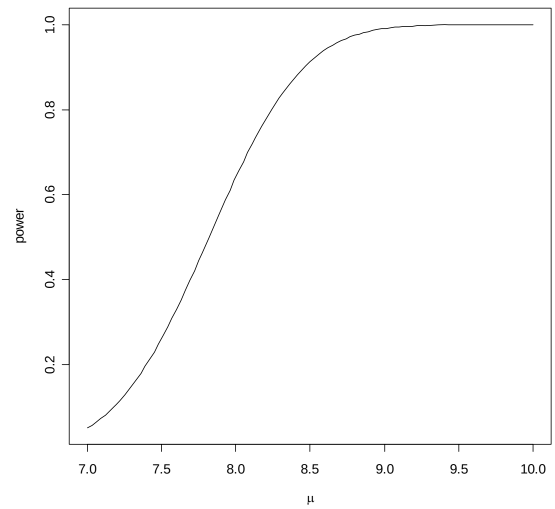
a. power(.1) = 1 – F_U(1.483) = .133 power(.4) = 1 – F_U(.884) = .609
power(.7) = 1 – F_U(.284) = .960 power(1) = 1 – F_U(–.316) = 1.



b. The power function is above.
c. Test 2 is a more powerful test.

10.91 Refer to Example 10.23 in the text. The hypotheses are H_0: μ = 7 vs. H_a: μ > 7.
a. The uniformly most powerful test is identically the Z–test from Section 10.3. The rejection region is: reject if $Z = \frac{\bar{Y}-7}{\sqrt{5/20}} > z_{.05} = 1.645$, or equivalently, reject if $\bar{Y} > 1.645\sqrt{.25} + 7 = 7.82$.
b. The power function is: power(μ) = $P(\bar{Y} > 7.82 \mid \mu) = P\left(Z > \frac{7.82-\mu}{\sqrt{5/20}}\right)$. Thus:

$$\text{power}(7.5) = P(\bar{Y} > 7.82 \mid 7.5) = P(Z > .64) = .2611.$$
$$\text{power}(8.0) = P(\bar{Y} > 7.82 \mid 8.0) = P(Z > -.36) = .6406.$$
$$\text{power}(8.5) = P(\bar{Y} > 7.82 \mid 8.5) = P(Z > -1.36) = .9131$$
$$\text{power}(9.0) = P(\bar{Y} > 7.82 \mid 9.0) = P(Z > -2.36) = .9909.$$



c. The power function is above.

10.92 Following Ex. 10.91, we require power(8) = $P(\overline{Y} > 7.82 \mid 8) = P\left(Z > \frac{7.82-8}{\sqrt{5/n}}\right)$ = .80. Thus, $\frac{7.82-8}{\sqrt{5/n}} = z_{.80} = -.84$. The solution is n = 108.89, or 109 observations must be taken.

10.93 Using the sample size formula from the end of Section 10.4, we have $n = \frac{(1.96+1.96)^2(25)}{(10-5)^2}$ = 15.3664, so 16 observations should be taken.

10.94 The most powerful test for H_0: $\sigma^2 = \sigma_0^2$ vs. H_a: $\sigma^2 = \sigma_1^2$, $\sigma_1^2 > \sigma_0^2$, is based on the likelihood ratio:

$$\frac{L(\sigma_0^2)}{L(\sigma_1^2)} = \left(\frac{\sigma_1}{\sigma_0}\right)^n \exp\left[-\left(\frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2}\sum_{i=1}^n (y_i - \mu)^2\right)\right] < k.$$

This simplifies to

$$T = \sum_{i=1}^n (y_i - \mu)^2 > \left[n \ln\left(\frac{\sigma_1}{\sigma_0}\right) - \ln k\right]\frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2} = c,$$

which is to say we should reject if the statistic T is large. To find a rejection region of size α, note that

$\frac{T}{\sigma_0^2} = \frac{\sum_{i=1}^n (Y_i - \mu)^2}{\sigma_0^2}$ has a chi–square distribution with n degrees of freedom. Thus, the most powerful test is equivalent to the chi–square test, and this test is UMP since the RR is the same for any $\sigma_1^2 > \sigma_0^2$.

10.95 **a.** To test H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, $\theta_0 < \theta_a$, the best test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^{12} \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_a}\right)\sum_{i=1}^4 y_i\right] < k.$$

This simplifies to

$$T = \sum_{i=1}^{4} y_i > \ln k \left(\frac{\theta_0}{\theta_a} \right)^{12} \left[\frac{1}{\theta_0} - \frac{1}{\theta_a} \right]^{-1} = c\,,$$

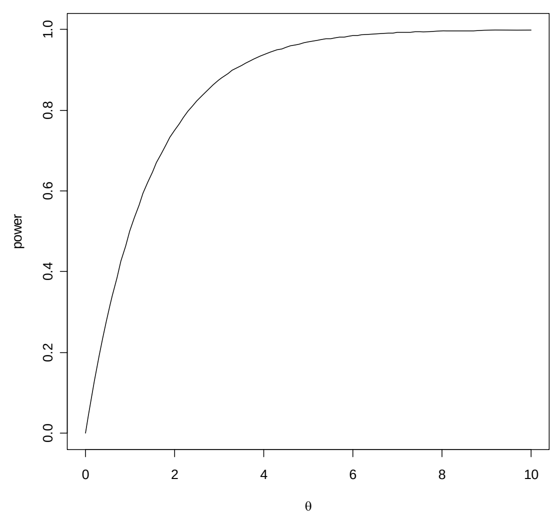
so H_0 should be rejected if T is large. Under H_0, Y has a gamma distribution with a shape parameter of 3 and scale parameter θ_0. Likewise, T is gamma with shape parameter of 12 and scale parameter θ_0, and $2T/\theta_0$ is chi–square with 24 degrees of freedom. The critical region can be written as

$$\frac{2T}{\theta_0} = \frac{2\sum_{i=1}^{4} Y_i}{\theta_0} > \frac{2c}{\theta_0} = c_1\,,$$

where c_1 will be chosen (from the chi–square distribution) so that the test is of size α.

b. Since the critical region doesn't depend on any specific $\theta_a < \theta_0$, the test is UMP.

10.96 **a.** The power function is given by power(θ) = $\int_{.5}^{1} \theta y^{\theta-1} dy = 1 - .5^{\theta}$. The power function is graphed below.



b. To test H_0: $\theta = 1$ vs. H_a: $\theta = \theta_a$, $1 < \theta_a$, the likelihood ratio is

$$\frac{L(1)}{L(\theta_a)} = \frac{1}{\theta_a y^{\theta_a - 1}} < k.$$

This simplifies to

$$y > \left(\frac{1}{\theta_a k} \right)^{\frac{1}{\theta_a - 1}} = c\,,$$

where c is chosen so that the test is of size α. This is given by

$$P(Y \geq c \mid \theta = 1) = \int_c^1 dy = 1 - c = \alpha\,,$$

so that $c = 1 - \alpha$. Since the RR does not depend on a specific $\theta_a > 1$, it is UMP.

10.97 Note that (N_1, N_2, N_3) is trinomial (multinomial with $k = 3$) with cell probabilities as given in the table.

a. The likelihood function is simply the probability mass function for the trinomial:

$$L(\theta) = \binom{n}{n_1\ n_2\ n_3}\theta^{2n_1}\left[2\theta(1-\theta)\right]^{n_2}(1-\theta)^{2n_3},\ 0 < \theta < 1,\ n = n_1 + n_2 + n_3.$$

b. Using part a, the best test for testing H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, $\theta_0 < \theta_a$, is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_0}{\theta_a}\right)^{2n_1+n_2}\left(\frac{1-\theta_0}{1-\theta_a}\right)^{n_2+2n_3} < k.$$

Since we have that $n_2 + 2n_3 = 2n - (2n_1 + n_2)$, the RR can be specified for certain values of $S = 2N_1 + N_2$. Specifically, the log–likelihood ratio is

$$s\ln\left(\frac{\theta_0}{\theta_a}\right) + (2n - s)\ln\left(\frac{1-\theta_0}{1-\theta_a}\right) < \ln k,$$

or equivalently

$$s > \left[\ln k - 2n\ln\left(\frac{1-\theta_0}{1-\theta_a}\right)\right] \times \left[\ln\left(\frac{\theta_0(1-\theta_a)}{\theta_a(1-\theta_0)}\right)\right]^{-1} = c\,.$$

So, the rejection region is given by $\{S = 2N_1 + N_2 > c\}$.

c. To find a size α rejection region, the distribution of (N_1, N_2, N_3) is specified and with $S = 2N_1 + N_2$, a null distribution for S can be found and a critical value specified such that $P(S \geq c \mid \theta_0) = \alpha$.

d. Since the RR doesn't depend on a specific $\theta_a > \theta_0$, it is a UMP test.

10.98 The density function that for the Weibull with shape parameter m and scale parameter θ.

a. The best test for testing H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, where $\theta_0 < \theta_a$, is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^n \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_a}\right)\sum_{i=1}^n y_i^m\right] < k,$$

This simplifies to

$$\sum_{i=1}^n y_i^m > -\left[\ln k + n\ln\left(\frac{\theta_0}{\theta_a}\right)\right] \times \left[\frac{1}{\theta_0} - \frac{1}{\theta_a}\right]^{-1} = c.$$

So, the RR has the form $\{T = \sum_{i=1}^m Y_i^m > c\}$, where c is chosen so the RR is of size α. To do so, note that the distribution of Y^m is exponential so that under H_0,

$$\frac{2T}{\theta_0} = \frac{2\sum_{i=1}^n Y_i^m}{\theta_0} > \frac{2c}{\theta_0}$$

is chi–square with $2n$ degrees of freedom. So, the critical value can be selected from the chi–square distribution and this does not depend on the specific $\theta_a > \theta_0$, so the test is UMP.

b. When H_0 is true, $T/50$ is chi–square with $2n$ degrees of freedom. Thus, $\chi^2_{.05}$ can be selected from this distribution so that the RR is $\{T/50 > \chi^2_{.05}\}$ and the test is of size $\alpha = .05$. If H_a is true, $T/200$ is chi–square with $2n$ degrees of freedom. Thus, we require

$$\beta = P(T/50 \le \chi^2_{.05} \mid \theta = 400) = P(T/200 \le \tfrac{1}{4}\chi^2_{.05} \mid \theta = 400) = P(\chi^2 \le \tfrac{1}{4}\chi^2_{.05}) = .05 .$$

Thus, we have that $\frac{1}{4}\chi^2_{.05} = \chi^2_{.95}$. From Table 6 in Appendix III, it is found that the degrees of freedom necessary for this equality is $12 = 2n$, so $n = 6$.

10.99 **a.** The best test is

$$\frac{L(\lambda_0)}{L(\lambda_a)} = \left(\frac{\lambda_0}{\lambda_a}\right)^T \exp[n(\lambda_a - \lambda_0)] < k,$$

where $T = \sum_{i=1}^{n} Y_i$. This simplifies to

$$T > \frac{\ln k - n(\lambda_a - \lambda_0)}{\ln(\lambda_0/\lambda_a)} = c,$$

and c is chosen so that the test is of size α.

b. Since under H_0 $T = \sum_{i=1}^{n} Y_i$ is Poisson with mean $n\lambda$, c can be selected such that

$$P(T > c \mid \lambda = \lambda_0) = \alpha.$$

c. Since this critical value does not depend on the specific $\lambda_a > \lambda_0$, so the test is UMP.

d. It is easily seen that the UMP test is: reject if $T < k'$.

10.100 Since **X** and **Y** are independent, the likelihood function is the product of all marginal mass function. The best test is given by

$$\frac{L_0}{L_1} = \frac{2^{\Sigma x_i + \Sigma y_i}\exp(-2m - 2n)}{\left(\frac{1}{2}\right)^{\Sigma x_i} 3^{\Sigma y_i}\exp(-m/2 - 3n)} = 4^{\Sigma x_i}\left(\tfrac{2}{3}\right)^{\Sigma y_i}\exp(-3m/2 + n) < k.$$

This simplifies to

$$(\ln 4)\sum_{i=1}^{m} x_i + \ln(2/3)\sum_{i=1}^{n} y_i < k',$$

and k' is chosen so that the test is of size α.

10.101 **a.** To test H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, where $\theta_a < \theta_0$, the best test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^n \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_a}\right)\sum_{i=1}^{n} y_i\right] < k.$$

Equivalently, this is

$$\sum_{i=1}^{n} y_i < \left[n\ln\left(\frac{\theta_0}{\theta_a}\right) + \ln k \right] \times \left[\frac{1}{\theta_a} - \frac{1}{\theta_0} \right]^{-1} = c\,,$$

and c is chosen so that the test is of size α (the chi–square distribution can be used – see Ex. 10.95).

b. Since the RR does not depend on a specific value of $\theta_a < \theta_0$, it is a UMP test.

10.102 **a.** The likelihood function is the product of the mass functions:

$$L(p) = p^{\Sigma y_i}(1-p)^{n-\Sigma y_i}\,.$$

i. It follows that the likelihood ratio is

$$\frac{L(p_0)}{L(p_a)} = \frac{p_0^{\Sigma y_i}(1-p_0)^{n-\Sigma y_i}}{p_a^{\Sigma y_i}(1-p_a)^{n-\Sigma y_i}} = \left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right)^{\Sigma y_i}\left(\frac{1-p_0}{1-p_a}\right)^{n}.$$

ii. Simplifying the above, the test rejects when

$$\sum_{i=1}^{n} y_i \ln\left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right) + n\ln\left(\frac{1-p_0}{1-p_a}\right) < \ln k.$$

Equivalently, this is

$$\sum_{i=1}^{n} y_i > \left[\ln k - n\ln\left(\frac{1-p_0}{1-p_a}\right)\right] \times \left[\ln\left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right)\right]^{-1} = c.$$

iii. The rejection region is of the form $\{\sum_{i=1}^{n} y_i > c\}$.

b. For a size α test, the critical value c is such that $P(\sum_{i=1}^{n} Y_i > c \mid p_0) = \alpha$. Under H_0, $\sum_{i=1}^{n} Y_i$ is binomial with parameters n and p_0.

c. Since the critical value can be specified without regard to a specific value of p_a, this is the UMP test.

10.103 Refer to Section 6.7 and 9.7 for this problem.

a. The likelihood function is $L(\theta) = \theta^{-n} I_{0,\theta}(y_{(n)})$. To test H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, where $\theta_a < \theta_0$, the best test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^{n} \frac{I_{0,\theta_0}(y_{(n)})}{I_{0,\theta_a}(y_{(n)})} < k.$$

So, the test only depends on the value of the largest order statistic $Y_{(n)}$, and the test rejects whenever $Y_{(n)}$ is small. The density function for $Y_{(n)}$ is $g_n(y) = ny^{n-1}\theta^{-n}$, for $0 \le y \le \theta$. For a size α test, select c such that

$$\alpha = P(Y_{(n)} < c \mid \theta = \theta_0) = \int_0^c ny^{n-1}\theta_0^{-n}dy = \frac{c^n}{\theta_0^n},$$

so $c = \theta_0\alpha^{1/n}$. So, the RR is $\{Y_{(n)} < \theta_0\alpha^{1/n}\}$.

b. Since the RR does not depend on the specific value of $\theta_a < \theta_0$, it is UMP.

10.104 Refer to Ex. 10.103.

a. As in Ex. 10.103, the test can be based on $Y_{(n)}$. In the case, the rejection region is of the form $\{Y_{(n)} > c\}$. For a size α test select c such that

$$\alpha = P(Y_{(n)} > c \mid \theta = \theta_0) = \int_c^{\theta_0} ny^{n-1}\theta_0^{-n}dy = 1 - \frac{c^n}{\theta_0^n},$$

so $c = \theta_0(1 - \alpha)^{1/n}$.

b. As in Ex. 10.103, the test is UMP.

c. It is not unique. Another interval for the RR can be selected so that it is of size α and the power is the same as in part a and independent of the interval. Example: choose the rejection region $C = (a,b) \cup (\theta_0, \infty)$, where $(a,b) \subset (0,\theta_0)$. Then,

$$\alpha = P(a < Y_{(n)} < b \mid \theta_0) = \frac{b^n - a^n}{\theta_0^n},$$

The power of this test is given by

$$P(a < Y_{(n)} < b \mid \theta_a) + P(Y_{(n)} > \theta_0 \mid \theta_a) = \frac{b^n - a^n}{\theta_a^n} + \frac{\theta_a^n - \theta_0^n}{\theta_a^n} = (\alpha - 1)\frac{\theta_0^n}{\theta_a^n} + 1,$$

which is independent of the interval (a, b) and has the same power as in part a.

10.105 The hypotheses are H_0: $\sigma^2 = \sigma_0^2$ vs. H_a: $\sigma^2 > \sigma_0^2$. The null hypothesis specifies $\Omega_0 = \{\sigma^2 : \sigma^2 = \sigma_0^2\}$, so in this restricted space the MLEs are $\hat{\mu} = \bar{y}$, σ_0^2. For the unrestricted space Ω, the MLEs are $\hat{\mu} = \bar{y}$, while

$$\hat{\sigma}^2 = \max\left[\sigma_0^2, \frac{1}{n}\sum_{i=1}^n (y_i - \bar{y})^2\right].$$

The likelihood ratio statistic is

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{2\sigma_0^2} + \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{2\hat{\sigma}^2}\right].$$

If $\hat{\sigma}^2 = \sigma_0^2$, $\lambda = 1$. If $\hat{\sigma}^2 > \sigma_0^2$,

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n\sigma_0^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{2\sigma_0^2} + \frac{n}{2}\right],$$

and H_0 is rejected when $\lambda \le k$. This test is a function of the chi–square test statistic $\chi^2 = (n-1)S^2/\sigma_0^2$ and since the function is monotonically decreasing function of χ^2, the test $\lambda \le k$ is equivalent to $\chi^2 \ge c$, where c is chosen so that the test is of size α.

10.106 The hypothesis of interest is H_0: $p_1 = p_2 = p_3 = p_4 = p$. The likelihood function is

$$L(\mathbf{p}) = \prod_{i=1}^{4} \binom{200}{y_i} p_i^{y_i} (1 - p_i)^{200-y_i}.$$

Under H_0, it is easy to verify that the MLE of p is $\hat{p} = \sum_{i=1}^{4} y_i / 800$. For the unrestricted space, $\hat{p}_i = y_i / 200$ for i = 1, 2, 3, 4. Then, the likelihood ratio statistic is

$$\lambda = \frac{\left(\frac{\Sigma y_i}{800}\right)^{\Sigma y_i} \left(1 - \frac{\Sigma y_i}{800}\right)^{800-\Sigma y_i}}{\prod_{i=1}^{4} \left(\frac{y_i}{200}\right)^{y_i} \left(1 - \frac{y_i}{200}\right)^{200-y_i}}.$$

Since the sample sizes are large, Theorem 10.2 can be applied so that $-2\ln\lambda$ is approximately distributed as chi–square with 3 degrees of freedom and we reject H_0 if $-2\ln\lambda > \chi^2_{.05} = 7.81$. For the data in this exercise, $y_1 = 76$, $y_2 = 53$, $y_3 = 59$, and $y_4 = 48$. Thus, $-2\ln\lambda = 10.54$ and we reject H_0: the fraction of voters favoring candidate A is not the sample in all four wards.

10.107 Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ denote the two samples. Under H_0, the quantity

$$V = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(Y_i - \overline{Y})^2}{\sigma_0^2} = \frac{(n-1)S_1^2 + (m-1)S_2^2}{\sigma_0^2}$$

has a chi–square distribution with $n + m - 2$ degrees of freedom. If H_a is true, then both S_1^2 and S_2^2 will tend to be larger than σ_0^2. Under H_0, the maximized likelihood is

$$L(\hat{\Omega}_0) = \frac{1}{(2\pi)^{n/2}} \frac{1}{\sigma_0^n} \exp\left(-\tfrac{1}{2}V\right).$$

In the unrestricted space, the likelihood is either maximized at σ_0 or σ_a. For the former, the likelihood ratio will be equal to 1. But, for $k < 1$, $\dfrac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} < k$ only if $\hat{\sigma} = \sigma_a$. In this case,

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\sigma_a}{\sigma_0}\right)^n \exp\left[-\tfrac{1}{2}V + \tfrac{1}{2}V\left(\tfrac{\sigma_0^2}{\sigma_a^2}\right)\right] = \left(\frac{\sigma_a}{\sigma_0}\right)^n \exp\left[-\tfrac{1}{2}V\left(1 - \tfrac{\sigma_0^2}{\sigma_a^2}\right)\right],$$

which is a decreasing function of V. Thus, we reject H_0 if V is too large, and the rejection region is $\{V > \chi^2_\alpha\}$.

10.108 The likelihood is the product of all $n = n_1 + n_2 + n_3$ normal densities:

$$L(\Theta) = \frac{1}{(2\pi)^n} \frac{1}{\sigma_1^{n_1}\sigma_2^{n_2}\sigma_3^{n_3}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n_1}\left(\frac{x_i - \mu_1}{\sigma_1}\right)^2 - \frac{1}{2}\sum_{i=1}^{n_2}\left(\frac{y_i - \mu_2}{\sigma_2}\right)^2 - \frac{1}{2}\sum_{i=1}^{n_3}\left(\frac{w_i - \mu_3}{\sigma_3}\right)^2\right\}$$

a. Under H_a (unrestricted), the MLEs for the parameters are:

$\hat{\mu}_1 = \overline{X}$, $\hat{\mu}_2 = \overline{Y}$, $\hat{\mu}_3 = \overline{W}$, $\hat{\sigma}_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(X_i - \overline{X})^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$ defined similarly.

Under H_0, $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ and the MLEs are

$$\hat{\mu}_1 = \overline{X},\ \hat{\mu}_2 = \overline{Y},\ \hat{\mu}_3 = \overline{W},\ \hat{\sigma}^2 = \frac{n_1\hat{\sigma}_1^2 + n_2\hat{\sigma}_2^2 + n_3\hat{\sigma}_3^2}{n}.$$

By defining the LRT, it is found to be equal to

$$\lambda = \frac{\left(\hat{\sigma}_1^2\right)^{n_1/2}\left(\hat{\sigma}_2^2\right)^{n_2/2}\left(\hat{\sigma}_3^2\right)^{n_3/2}}{\left(\hat{\sigma}^2\right)^{n/2}}.$$

b. For large values of n_1, n_2, and n_3, the quantity $-2\ln\lambda$ is approximately chi–square with 3–1=2 degrees of freedom. So, the rejection region is: $-2\ln\lambda > \chi^2_{.05} = 5.99$.

10.109 The likelihood function is $L(\Theta) = \frac{1}{\theta_1^m\theta_2^n}\exp\left[-\left(\sum_{i=1}^m x_i/\theta_1 + \sum_{i=1}^n y_i/\theta_2\right)\right]$.

a. Under H_a (unrestricted), the MLEs for the parameters are:

$$\hat{\theta}_1 = \overline{X},\ \hat{\theta}_2 = \overline{Y}.$$

Under H_0, $\theta_1 = \theta_2 = \theta$ and the MLE is

$$\hat{\theta} = (m\overline{X} + n\overline{Y})/(m+n).$$

By defining the LRT, it is found to be equal to

$$\lambda = \frac{\overline{X}^m\overline{Y}^n}{\left(\frac{m\overline{X}+n\overline{Y}}{m+n}\right)^{m+n}}$$

b. Since $2\sum_{i=1}^m X_i/\theta_1$ is chi–square with $2m$ degrees of freedom and $2\sum_{i=1}^n Y_i/\theta_2$ is chi–square with $2n$ degrees of freedom, the distribution of the quantity under H_0

$$F = \frac{\dfrac{\left(2\sum_{i=1}^m X_i/\theta\right)}{2m}}{\dfrac{\left(2\sum_{i=1}^n Y_i/\theta\right)}{2n}} = \frac{\overline{X}}{\overline{Y}}$$

has an F–distribution with $2m$ numerator and $2n$ denominator degrees of freedom. This test can be seen to be equivalent to the LRT in part a by writing

$$\lambda = \frac{\overline{X}^m\overline{Y}^n}{\left(\frac{m\overline{X}+n\overline{Y}}{m+n}\right)^{m+n}} = \left[\frac{m\overline{X}+n\overline{Y}}{\overline{X}(m+n)}\right]^{-m}\left[\frac{m\overline{X}+n\overline{Y}}{\overline{Y}(m+n)}\right]^{-n} = \left[\frac{m}{m+n} + \frac{n}{F(m+n)}\right]^{-m}\left[\frac{m}{m+n}F + \frac{n}{m+n}\right]^{-n}.$$

So, λ is small if F is too large or too small. Thus, the rejection region is equivalent to $F > c_1$ and $F < c_2$, where c_1 and c_2 are chosen so that the test is of size α.

10.110 This is easily proven by using Theorem 9.4: write the likelihood function as a function of the sufficient statistic, so therefore the LRT must also only be a function of the sufficient statistic.

10.111 **a.** Under H_0, the likelihood is maximized at θ_0. Under the alternative (unrestricted) hypothesis, the likelihood is maximized at either θ_0 or θ_a. Thus, $L(\hat{\Omega}_0) = L(\theta_0)$ and $L(\hat{\Omega}) = \max\{L(\theta_0), L(\theta_a)\}$. Thus,

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{L(\theta_0)}{\max\{L(\theta_0), L(\theta_a)\}} = \frac{1}{\max\{1,\ L(\theta_a)/L(\theta_0)\}}.$$

b. Since $\frac{1}{\max\{1,\ L(\theta_a)/L(\theta_0)\}} = \min\{1,\ L(\theta_0)/L(\theta_a)\}$, we have $\lambda < k < 1$ if and only if

$$L(\theta_0)/L(\theta_a) < k.$$

c. The results are consistent with the Neyman–Pearson lemma.

10.112 Denote the samples as $X_1, \ldots, X_{n_1}$, and $Y_1, \ldots, Y_{n_2}$, where $n = n_1 + n_2$.

Under H_a (unrestricted), the MLEs for the parameters are:

$$\hat{\mu}_1 = \overline{X},\ \hat{\mu}_2 = \overline{Y},\ \hat{\sigma}^2 = \tfrac{1}{n}\left(\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2\right).$$

Under H_0, $\mu_1 = \mu_2 = \mu$ and the MLEs are

$$\hat{\mu} = \tfrac{n_1\overline{X} + n_2\overline{Y}}{n},\ \hat{\sigma}_0^2 = \tfrac{1}{n}\left(\sum_{i=1}^{n_1}(X_i - \hat{\mu})^2 + \sum_{i=1}^{n_2}(Y_i - \hat{\mu})^2\right).$$

By defining the LRT, it is found to be equal to

$$\lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \le k \text{, or equivalently reject if } \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right) \ge k'.$$

Now, write

$$\sum_{i=1}^{n_1}(X_i - \hat{\mu})^2 = \sum_{i=1}^{n_1}(X_i - \overline{X} + \overline{X} - \hat{\mu})^2 = \sum_{i=1}^{n_1}(X_i - \overline{X})^2 + n_1(\overline{X} - \hat{\mu})^2,$$

$$\sum_{i=1}^{n_2}(Y_i - \hat{\mu})^2 = \sum_{i=1}^{n_2}(Y_i - \overline{Y} + \overline{Y} - \hat{\mu})^2 = \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + n_2(\overline{Y} - \hat{\mu})^2,$$

and since $\hat{\mu} = \frac{n_1}{n}\overline{X} + \frac{n_2}{n}\overline{Y}$, and alternative expression for $\hat{\sigma}_0^2$ is

$$\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + \tfrac{n_1 n_2}{n}(\overline{X} - \overline{Y})^2.$$

Thus, the LRT rejects for large values of

$$1 + \tfrac{n_1 n_2}{n}\left(\frac{(\overline{X} - \overline{Y})^2}{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}\right).$$

Now, we are only concerned with $\mu_1 > \mu_2$ in H_a, so we could only reject if $\overline{X} - \overline{Y} > 0$.

Thus, the test is equivalent to rejecting if $\dfrac{\overline{X} - \overline{Y}}{\sqrt{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}}$ is large.

This is equivalent to the two–sample t test statistic (σ^2 unknown) except for the constants that do not depend on the data.

10.113 Following Ex. 10.112, the LRT rejects for large values of

$$1 + \tfrac{n_1 n_2}{n}\left(\frac{(\overline{X} - \overline{Y})^2}{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}\right).$$

Equivalently, the test rejects for large values of

$$\frac{|\overline{X}-\overline{Y}|}{\sqrt{\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2}}.$$

This is equivalent to the two–sample t test statistic (σ^2 unknown) except for the constants that do not depend on the data.

10.114 Using the sample notation $Y_{11}, \ldots, Y_{1n_1}, Y_{21}, \ldots, Y_{2n_2}, Y_{31}, \ldots, Y_{3n_3}$, with $n = n_1 + n_2 + n_3$, we have that under H_a (unrestricted hypothesis), the MLEs for the parameters are:

$$\hat{\mu}_1 = \overline{Y}_1,\ \hat{\mu}_2 = \overline{Y}_2,\ \hat{\mu}_3 = \overline{Y}_3,\ \hat{\sigma}^2 = \tfrac{1}{n}\left(\sum_{i=1}^{3}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2\right).$$

Under H_0, $\mu_1 = \mu_2 = \mu_3 = \mu$ so the MLEs are

$$\hat{\mu} = \tfrac{1}{n}\sum_{i=1}^{3}\sum_{j=1}^{n_i}Y_{ij} = \tfrac{n_1\overline{Y}_1+n_2\overline{Y}_2+n_3\overline{Y}_3}{n},\quad \hat{\sigma}_0^2 = \tfrac{1}{n}\left(\sum_{i=1}^{3}\sum_{j=1}^{n_i}(Y_{ij}-\hat{\mu})^2\right).$$

Similar to Ex. 10.112, ny defining the LRT, it is found to be equal to

$$\lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \le k \text{, or equivalently reject if } \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right) \ge k'.$$

In order to show that this test is equivalent to and exact F test, we refer to results and notation given in Section 13.3 of the text. In particular,

$n\hat{\sigma}^2$ = SSE

$n\hat{\sigma}_0^2$ = TSS = SST + SSE

Then, we have that the LRT rejects when

$$\frac{\text{TSS}}{\text{SSE}} = \frac{\text{SSE}+\text{SST}}{\text{SSE}} = 1+\frac{\text{SST}}{\text{SSE}} = 1+\frac{\text{MST}}{\text{MSE}}\tfrac{2}{n-3}+1+F\tfrac{2}{n-3} \ge k',$$

where the statistic $F = \dfrac{\text{MST}}{\text{MSE}} = \dfrac{\text{SST/2}}{\text{SSE}/(n\text{-}3)}$ has an F–distribution with 2 numerator and n–3 denominator degrees of freedom under H_0. The LRT rejects when the statistic F is large and so the tests are equivalent,

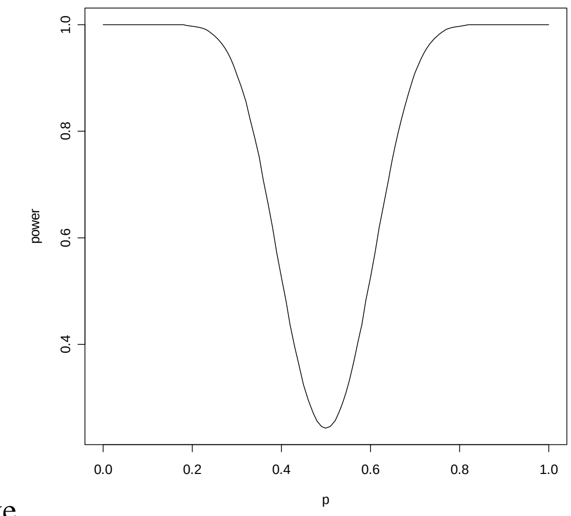
10.115 **a.** True
b. False: H_0 is not a statement regarding a random quantity.
c. False: "large" is a relative quantity
d. True
e. False: power is computed for specific values in H_a
f. False: it must be true that p–value $\le \alpha$
g. False: the UMP test has the highest power against all other α–level tests.
h. False: it always holds that $\lambda \le 1$.
i. True.

10.116 From Ex. 10.6, we have that

$$\text{power}(p) = 1 - \beta(p) = 1 - P(|Y - 18| \le 3 \mid p) = 1 - \text{P}(15 \le \text{Y} \le 21 \mid p).$$

Thus,

power(.2) = .9975 power(.3) = .9084 power(.4) = .5266
power(.5) = .2430
power(.6) = .9975 power(.7) = .9084 power(.8) = .5266



A graph of the power function is above.

10.117 **a.** The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1 = mean nitrogen density for chemical compounds and μ_2 = mean nitrogen density for air. Then, $s_p^2 = \frac{9(.00131)^2 + 8(.000574)^2}{17} = .000001064$ and $|t| = 22.17$ with 17 degrees of freedom. The p–value is far less than $2(.005) = .01$ so H_0 should be rejected.

b. The 95% CI for $\mu_1 - \mu_2$ is (–.01151, –.00951).
c. Since the CI do not contain 0, there is evidence that the mean densities are different.
d. The two approaches agree.

10.118 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 < 0$, where μ_1 = mean alcohol blood level for sea level and μ_2 = mean alcohol blood level for 12,000 feet. The sample statistics are $\bar{y}_1 = .10$, $s_1 = .0219$, $\bar{y}_2 = .1383$, $s_2 = .0232$. The computed value of the test statistic is $t = -2.945$ and with 10 degrees of freedom, $-t_{.10} = -1.383$ so H_0 should be rejected.

10.119 **a.** The hypotheses are H_0: $p = .20$, H_a: $p > .20$.
b. Let Y = # who prefer brand A. The significance level is
$\alpha = P(Y \geq 92 \mid p = .20) = P(Y > 91.5 \mid p = .20) \approx P(Z > \frac{91.5-80}{8}) = P(Z > 1.44) = .0749$.

10.120 Let μ = mean daily chemical production.
a. H_0: $\mu = 1100$, H_a: $\mu < 1100$.
b. With .05 significance level, we can reject H_0 if $Z < -1.645$.
c. For this large sample test, $Z = -1.90$ and we reject H_0: there is evidence that suggests there has been a drop in mean daily production.

10.121 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean breaking distances. For this large–sample test, the computed test statistic is $|z| = \frac{|118-109|}{\sqrt{\frac{102}{64}+\frac{87}{64}}} = 5.24$. Since p–value ≈ $2P(Z > 5.24)$ is approximately 0, we can reject the null hypothesis: the mean braking distances are different.

10.122 **a.** To test $H_0\!: \sigma_1^2 = \sigma_2^2$ vs. $H_a\!: \sigma_1^2 > \sigma_2^2$, where σ_1^2, σ_2^2 represent the population variances for the two lines, the test statistic is $F = (92{,}000)/(37{,}000) = 2.486$ with 49 numerator and 49 denominator degrees of freedom. So, with $F_{.05} = 1.607$ we can reject the null hypothesis.

b. p–value = $P(F > 2.486) = .0009$
Using R:

```
> 1-pf(2.486,49,49)
[1] 0.0009072082
```

10.123 **a.** Our test is $H_0\!: \sigma_1^2 = \sigma_2^2$ vs. $H_a\!: \sigma_1^2 \neq \sigma_2^2$, where σ_1^2, σ_2^2 represent the population variances for the two suppliers. The computed test statistic is $F = (.273)/(.094) = 2.904$ with 9 numerator and 9 denominator degrees of freedom. With $\alpha = .05$, $F_{.05} = 3.18$ so H_0 is not rejected: we cannot conclude that the variances are different.

b. The 90% CI is given by $\left(\frac{9(.094)}{16.919}, \frac{9(.094)}{3.32511}\right) = (.050, .254)$. We are 90% confident that the true variance for Supplier B is between .050 and .254.

10.124 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean strengths for the two materials. Then, $s_p^2 = .0033$ and $t = \frac{1.237-.978}{\sqrt{.0033\left(\frac{2}{9}\right)}} = 9.568$ with 17 degrees of freedom. With $\alpha = .10$, the critical value is $t_{.05} = 1.746$ and so H_0 is rejected.

10.125 **a.** The hypotheses are H_0: $\mu_A - \mu_B = 0$ vs. H_a: $\mu_A - \mu_B \neq 0$, where μ_A, μ_B are the mean efficiencies for the two types of heaters. The two sample means are 73.125, 77.667, and $s_p^2 = 10.017$. The computed test statistic is $\frac{73.125-77.667}{\sqrt{10.017\left(\frac{1}{8}+\frac{1}{6}\right)}} = -2.657$ with 12 degrees of freedom. Since p–value = $2P(T > 2.657)$, we obtain $.02 < p\text{–value} < .05$ from Table 5 in Appendix III.

b. The 90% CI for $\mu_A - \mu_B$ is

$$73.125 - 77.667 \pm 1.782\sqrt{10.017\left(\tfrac{1}{8}+\tfrac{1}{6}\right)} = -4.542 \pm 3.046 \text{ or } (-7.588, -1.496).$$

Thus, we are 90% confident that the difference in mean efficiencies is between –7.588 and –1.496.

10.126 **a.** $SE(\hat{\theta}) = \sqrt{V(\hat{\theta})} = \sqrt{a_1^2 V(\overline{X}) + a_2^2 V(\overline{Y}) + a_3^2 V(\overline{W})} = \sigma\sqrt{\frac{a_1^2}{n_1} + \frac{a_2^2}{n_1} + \frac{a_3^2}{n_3}}$.

b. Since $\hat{\theta}$ is a linear combination of normal random variables, $\hat{\theta}$ is normally distributed with mean θ and standard deviation given in part **a**.
c. The quantity $(n_1 + n_2 + n_3)S_p^2/\sigma^2$ is chi–square with $n_1+n_2+n_3 - 3$ degrees of freedom and by Definition 7.2, T has a t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom.

d. A 100(1 – α)% CI for θ is $\hat{\theta} \pm t_{\alpha/2} s_p \sqrt{\frac{a_1^2}{n_1} + \frac{a_2^2}{n_1} + \frac{a_3^2}{n_3}}$, where $t_{\alpha/2}$ is the upper–α/2 critical value from the t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom.

e. Under H_0, the quantity $t = \dfrac{(\hat{\theta} - \theta_0)}{s_p \sqrt{\frac{a_1^2}{n_1} + \frac{a_2^2}{n_1} + \frac{a_3^2}{n_3}}}$ has a t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom. Thus, the rejection region is: $|t| > t_{\alpha/2}$.

10.127 Let $P = X + Y - W$. Then, P has a normal distribution with mean $\mu_1 + \mu_2 - \mu_3$ and variance $(1 + a + b)\sigma^2$. Further, $\overline{P} = \overline{X} + \overline{Y} - \overline{W}$ is normal with mean $\mu_1 + \mu_2 - \mu_3$ and variance $(1 + a + b)\sigma^2/n$. Therefore,

$$Z = \frac{\overline{P} - (\mu_1 + \mu_2 - \mu_3)}{\sigma\sqrt{(1 + a + b)/n}}$$

is standard normal. Next, the quantities

$$\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2},\quad \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{a\sigma^2},\quad \frac{\sum_{i=1}^{n}(W_i - \overline{W})^2}{b\sigma^2}$$

have independent chi–square distributions, each with $n - 1$ degrees of freedom. So, their sum is chi–square with $3n - 3$ degrees of freedom. Therefore, by Definition 7.2, we can build a random variable that follows a t–distribution (under H_0) by

$$T = \frac{\overline{P} - k}{S_p\sqrt{(1 + a + b)/n}},$$

where $S_P^2 = \left(\sum_{i=1}^{n}(X_i - \overline{X})^2 + \frac{1}{a}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 + \frac{1}{b}\sum_{i=1}^{n}(W_i - \overline{W})^2\right)\Big/(3n - 3)$. For the test, we reject if $|t| > t_{.025}$, where $t_{.025}$ is the upper .024 critical value from the t–distribution with $3n - 3$ degrees of freedom.

10.128 The point of this exercise is to perform a "two–sample" test for means, but information will be garnered from three samples – that is, the common variance will be estimated using three samples. From Section 10.3, we have the standard normal quantity

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

As in Ex. 10.127, $\left(\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + \sum_{i=1}^{n_3}(W_i - \overline{W})^2\right)\Big/\sigma^2$ has a chi–square distribution with $n_1+n_2+n_3 - 3$ degrees of freedom. So, define the statistic

$$S_P^2 = \left(\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + \sum_{i=1}^{n_3}(W_i - \overline{W})^2\right)\Big/(n_1 + n_2 + n_3 - 3)$$

and thus the quantity $T = \dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ has a t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom.

For the data given in this exercise, we have H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$ and with $s_P = 10$, the computed test statistic is $|t| = \frac{|60-50|}{10\sqrt{\frac{2}{10}}} = 2.326$ with 27 degrees of freedom. Since $t_{.025} = 2.052$, the null hypothesis is rejected.

10.129 The likelihood function is $L(\Theta) = \theta_1^{-n}\exp[-\sum_{i=1}^n (y_i - \theta_2)/\theta_1]$. The MLE for θ_2 is $\hat{\theta}_2 = Y_{(1)}$. To find the MLE of θ_1, we maximize the log–likelihood function to obtain $\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^n (Y_i - \hat{\theta}_2)$. Under H_0, the MLEs for θ_1 and θ_2 are (respectively) $\theta_{1,0}$ and $\hat{\theta}_2 = Y_{(1)}$ as before. Thus, the LRT is

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\hat{\theta}_1}{\theta_{1,0}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (y_i - y_{(1)})}{\theta_{1,0}} + \frac{\sum_{i=1}^n (y_i - y_{(1)})}{\hat{\theta}_1}\right]$$

$$= \left(\frac{\sum_{i=1}^n (y_i - y_{(1)})}{n\theta_{1,0}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (y_i - y_{(1)})}{\theta_{1,0}} + n\right].$$

Values of $\lambda \leq k$ reject the null hypothesis.

10.130 Following Ex. 10.129, the MLEs are $\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^n (Y_i - \hat{\theta}_2)$ and $\hat{\theta}_2 = Y_{(1)}$. Under H_0, the MLEs for θ_2 and θ_1 are (respectively) $\theta_{2,0}$ and $\hat{\theta}_{1,0} = \frac{1}{n}\sum_{i=1}^n (Y_i - \theta_{2,0})$. Thus, the LRT is given by

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\hat{\theta}_1}{\hat{\theta}_{1,0}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (y_i - \theta_{2,0})}{\hat{\theta}_{1,0}} + \frac{\sum_{i=1}^n (y_i - y_{(1)})}{\hat{\theta}_1}\right] = \left[\frac{\sum_{i=1}^n (y_i - y_{(1)})}{\sum_{i=1}^n (y_i - \theta_{2,0})}\right]^n.$$

Values of $\lambda \leq k$ reject the null hypothesis.