MATH 323

Solutions to Assignment No. 5

9.62 Since $\mu_1' = \lambda$, we equate $\overline{Y} = \Sigma \frac{Y_i}{n}$ to μ_1' and obtain $\hat{\lambda} = \overline{Y}$.

9.64 Similar to Exercise 9.61, with μ unknown. The two equations to be solved are

$$\overline{Y} = \mu$$

and

$$\frac{\Sigma Y_i^2}{n} = \mu_2' = \hat{\sigma}^2 + \hat{\mu}^2.$$

So that

$$\hat{\mu} = \overline{Y}$$

and

$$\hat{\sigma}^2 = \frac{\Sigma Y_i^2}{n} - \overline{Y}^2 = \frac{1}{n}\left[\Sigma (Y_i - \overline{Y})^2\right].$$

9.66 a. Calculate

$$\mu_1' = E(Y) = \frac{2}{\theta^2}\int_0^\theta (\theta y - y^2)\,dy = \frac{2}{\theta^2}\left[\frac{\theta y^2}{2} - \frac{y^3}{3}\right]_0^\theta = \frac{\theta}{3}$$

Equating sample and population moments, we have $\frac{\hat{\theta}}{3} = \overline{Y}$, or $\hat{\theta} = 3\overline{Y}$.

9.68 For a single observation y, the first sample moment is $m_1' = Y$, and since Y has a geometric distribution, $E(Y) = \mu_1' = \frac{1}{p}$. Hence the moment estimator of p is $\hat{p} = \frac{1}{Y}$.

9.70 $E(Y) = \int_0^3 \frac{\alpha y^\alpha}{3^\alpha}\,dy = \left(\frac{\alpha}{3^\alpha}\right)\left(\frac{y^{\alpha+1}}{\alpha+1}\right)\Big]_0^3 = \frac{3\alpha}{\alpha+1} = \overline{Y}.$

$\Rightarrow 3\alpha = \alpha\overline{y} + \overline{y}$

$\Rightarrow (3 - \overline{y})\,\alpha = \overline{y}$

$\Rightarrow \hat{\alpha} = \frac{\overline{Y}}{3 - \overline{Y}}.$

2

9.72 a. The likelihood function is

$$L = \prod_{i=1}^{n} \frac{\lambda^{y_i} e^{-\lambda}}{y!} = \frac{\lambda^{\Sigma y_i} e^{-n\lambda}}{\prod_{i=1}^{n} y_i!}$$

and

$$\ln L = (\Sigma y_i) \ln \lambda - n\lambda - \Sigma \ln y_i!.$$

so that $\left(\frac{d}{d\lambda}\right) [\ln L] = \left(\Sigma \frac{y_i}{\lambda}\right) - n$. Equating the derivative to 0, we obtain

$$\frac{\Sigma y_i}{\hat{\lambda}} - n = 0$$

or

$$\hat{\lambda} = \frac{\Sigma Y_i}{n} = \overline{Y}.$$

b. Recalling that $E(Y_i) = \lambda$ and $V(Y_i) = \lambda$, we obtain

$$E(\hat{\lambda}) = \frac{\sum_{i=1}^{n} E(Y_i)}{n} = \lambda$$

and

$$V(\hat{\lambda}) = \frac{\sum_{i=1}^{n} V(Y_i)}{n^2} = \frac{\lambda}{n}.$$

c. Since $E(Y_i) = \lambda$ and $V(Y_i) = \lambda < \infty$, the law of large numbers applies and we conclude that $\hat{\lambda}$ converges in probability to λ. Hence $\hat{\lambda}$ is consistent for λ.

d. The MLE of λ was found in part **a** to be $\hat{\lambda} = \overline{Y}$. Then, the MLE for $e^{-\lambda}$ is $e^{-\overline{Y}}$.

9.74

b. Consider $\ln L = n \ln r - n \ln \theta + (r-1) \sum_{i=1}^{n} \ln y_i - \sum_{i=1}^{n} \frac{y_i^r}{\theta}$ and $\frac{d}{d\theta} \ln L = \frac{-n}{\theta} + \frac{\Sigma y_i^r}{\theta^2}$.

Equating the derivative to 0, the estimator is obtained.

$$\frac{-n}{\hat{\theta}} + \frac{\Sigma y_i^r}{\hat{\theta}^2} = 0 \qquad \text{or} \qquad -n\hat{\theta} + \Sigma y_i^r = 0 \qquad \text{or} \qquad \hat{\theta} = \frac{\Sigma Y_i^r}{n}.$$

9.76 a. As this exercise is a special case of exercise 9.77 **a** (with $\alpha = 2$) we will refer to its results.

$$\hat{\theta} = \left(\frac{\overline{Y}}{2}\right) = \frac{378}{3(2)} = 63.$$

b. From Exercise 9.69 **b**,

$$E(\hat{\theta}) = \theta \qquad V(\hat{\theta}) = \frac{\theta^2}{n\alpha} = \frac{\theta^2}{3(2)} = \frac{\theta^2}{6}$$

c. The bound on the error of estimation is

$$2\sqrt{V(\hat{\theta})} = 2\sqrt{\frac{\theta^2}{6}} = 2\sqrt{\frac{(130)^2}{6}} = 106.14$$

d. The variance of Y is $2\theta^2$. The MLE of θ was found in part a to be $\hat{\theta} = 63$. Therefore, the MLE for the variance is $2(63)^2 = 7938$.

9.78 The likelihood function is

$$L=\prod_{i=1}^{m}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\tfrac{1}{2}\left(\tfrac{x_i-\mu_1}{\sigma}\right)^2\right]\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\theta}\exp\left[-\tfrac{1}{2}\left(\tfrac{y_i-\mu_2}{\sigma}\right)^2\right]$$

$$=\frac{1}{(2\pi)^{(m+n)/2}\sigma^{m+n}}\exp\left\{-\tfrac{1}{2}\left[\sum_{i=1}^{m}\left(\tfrac{x_i-\mu_1}{\sigma}\right)^2+\sum_{i=1}^{n}\left(\tfrac{y_i-\mu_2}{\sigma}\right)^2\right]\right\}$$

and

$$\ln L=\ln K-(m+n)\ln\sigma-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{i=1}^{n}(y_i-\mu_2)^2\right]$$

Then

$$\frac{d}{d\sigma}\ln L=\frac{-(m+n)}{\sigma}+\frac{1}{\sigma^3}\left[\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{i=1}^{n}(y_i-\mu_2)^2\right]$$

Setting the derivative equal to 0 and solving for $\hat{\sigma}$, we have

$$\frac{m+n}{\hat{\sigma}}=\frac{1}{\hat{\sigma}^3}\left[\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{i=1}^{n}(y_i-\mu_2)^2\right]$$

or

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{n}(X_i-\mu_1)^2+\sum_{i=1}^{n}(Y_i-\mu_2)^2}{m+n}.$$

Since μ_1 and μ_2 are unknown, their maximum likelihood estimates must be obtained.

$$\frac{d}{d\mu_1}\ln L=\frac{\sum_{i=1}^{m}(x_i-\mu_1)}{\sigma^2}\quad\text{and}\quad\frac{d}{d\mu_2}\ln L=\frac{\sum_{i=1}^{n}(y_i-\mu_2)}{\sigma^2}$$

and, as in Example 9.15 in the text, $\hat{\mu}_1=\overline{X}$ and $\hat{\mu}_2=\overline{Y}$. Thus,

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{m}(X_i-\overline{X})^2+\sum_{i=1}^{n}(Y_i-\overline{Y})^2}{m+n}$$

9.80 The likelihood function is

$$L=\prod_{i=1}^{n}(\theta+1)y_i^{\theta}=(\theta+1)^n\prod_{i=1}^{n}y_i^{\theta}\quad\text{and}\quad\ln L=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln y_i$$

so that

$$\frac{d}{d\theta}\ln L=\frac{n}{\theta+1}+\Sigma\ln y_i$$

Equating the derivative to 0, we can find the maximum likelihood estimator,

$$\frac{n}{\hat{\theta}+1}+\Sigma\ln y_i=0\quad\text{or}\quad n+(\hat{\theta}+1)\,\Sigma\ln y_i=0\quad\text{or}\quad\hat{\theta}=\frac{-n}{\Sigma\ln Y_i}-1$$

9.84 a. Recall from exercise 9.44 $L(y_1,\ldots,y_n|\theta)=\begin{cases}\frac{3^n}{\theta^{3n}}\left(\prod_{i=1}^{n}y_i\right) & \text{for } y_{(n)}\le\theta\\ 0 & \text{Otherwise}\end{cases}$

That is $L(y_1,\ldots,y_n|\theta)=\frac{c}{\theta^{3n}}$ where $c=3^n\left(\prod_{i=1}^{n}y_i\right)$ when $\theta\ge y_{(n)}$ (and 0 elsewhere). This is a decreasing function of θ. Therefore it will it will obtain its maximum where θ is smallest, i.e. when $\theta=Y_{(n)}$. Thus the MLE of θ is $Y_{(n)}$.

10.2 The test statistic Y has a binomial distribution with $n = 20$ and p.

a. A Type I error occurs if the experimenter concluded that the drug dosage level induces sleep in less than 80% of the people suffering from insomnia when, in fact, drug dosage level does induce sleep in 80% of insomniacs.

b. $\alpha = P(\text{reject } H_0 | H_0 \text{ true}) = P(Y \leq 12 | p = .8) = .032$, using Table 1, Appendix III.

c. A Type II error would occur if the experimenter concluded that the drug dosage level induces sleep in 80% of the people suffering from insomnia when, in fact, fewer than 80% experience relief.

d. If $p = .6$,

$$\beta = P(\text{accept } H_0 | H_0 \text{ false}) = P(Y > 12 | p = .6) = 1 - P(Y \leq 12 p = .6) = 1 - .584 = .416$$

e. If $p = .4$, then

$$\beta = P(Y > 12 | p = .4) = 1 - P(Y \leq 12 | p = .4) = 1 - .979 = .021.$$

10.4 **a.** A Type I error occurs if we conclude that the proportion of ledger sheets with errors is larger than .05 when, in fact, the proportion is .05.

b. By the scheme being used, we will reject for the following situations:
(NOTE: NE = no error, E = error)

Sheet 1	Sheet 2	Sheet 3
NE	NE	.
NE	E	NE
E	NE	NE
E	E	NE

Thus, $\alpha = (.95)^2 + 2(.05)(.95)^2 + (.05)^2(.95) = .9025 + .09025 + .002375 = .995125.$

c. A Type II error occurs if we conclude that the proportion of ledger sheets with errors is .05 when, in fact, the proportion is larger than .05.

d. $\beta = P(\text{accept } H_0 \text{ when } H_a \text{ is true}) = P(\text{accepting } H_0 | p = p_a)$
$= 2p_a^2(1 - p_a) + p_a^3$. Since we reject if we observe E,E,E or NE,E,E or E,NE,E,

10.6 **a.** Let X_1 and X_2 both be binomial with $n = 15$ and $p = 0.10$. By definition,

$$\begin{aligned}
\alpha &= P(\text{type I error}) = P(\text{reject } H_0 | H_0 \text{is true}) \\
&= P(\text{reject } H_0 \text{ in stage } 1 | H_0 \text{ is true}) + P(\text{reject } H_0 \text{ in stage } 2 | H_0 \text{is true}) \\
&= P(X_1 \geq 4) + P(X_1 + X_2 \geq 6 \text{ and } X_1 \leq 3) \\
&= P(X_1 \geq 4) + \sum_{i=0}^{3} P(X_1 + X_2 \geq 6 \text{ and } X_1 = i) \\
&= P(X_1 \geq 4) + \sum_{i=0}^{3} P(X_2 \geq 6 - i)\, P(X_1 = i) \\
&= 1 - P(X_1 \leq 3) + \sum_{i=0}^{3} [1 - P(X_2 \leq 5 - i)]\,[P(X_1 \leq i) - P(X_1 \leq i - 1)] \\
&= 1 - 0.944 + (1 - 0.998)(0.206) + \cdots + (1 - 0.816)(0.944 - 0.816). \\
&= 0.0994
\end{aligned}$$

b. This is similar to part a. with $p = 0.30$.

c. Let X_1 and X_2 be binomial with $n = 15$ and $p = 0.30$. By definition,

$$\begin{aligned}
\beta &= P(\text{type II error}) = P(\text{accept } H_0 | p = 0.30) \\
&= \sum_{i=0}^{3} P(X_1 = i \text{ and } X_1 + X_2 \leq 5) \\
&= \sum_{i=0}^{3} P(X_2 \leq 5 - i) P(X_1 = i) \\
&= 0.068
\end{aligned}$$

10.8 The parameter of interest is μ, the average daily wage of workers in a given company. The objective is to determine whether this company pays inferior wages in comparison to the total industry. Thus the hypothesis to be tested is

$$H_0\colon\ \mu = 13.20 \qquad \text{vs.} \qquad H_a\colon\ \mu < 13.20$$

The best estimator for μ is the sample average, $\overline{y} = 12.20$. The test statistic is

$$Z = \frac{\overline{Y}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

which represents the distance (measured in units of standard deviation) from $\overline{Y}$ to the hypothesized mean μ. Calculating the value of the test statistic using the information contained in the sample, we have

$$z = \frac{\overline{y}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{12.20-13.20}{\frac{2.50}{\sqrt{40}}} = \frac{-\sqrt{40}}{2.5} = -2.53$$

The critical value of Z that separates the rejection and non rejection regions will be a value (denoted by z_0) such that $P(Z < z_0) = .01$. That is, $z_0 = -2.326$ (see Figure 10.2). The null hypothesis will be rejected if $z < -2.326$. Note that the observed value of the test statistic falls in the rejection region. Thus the conclusion is to reject the null hypothesis. There is evidence to indicate that this company is paying inferior wages.

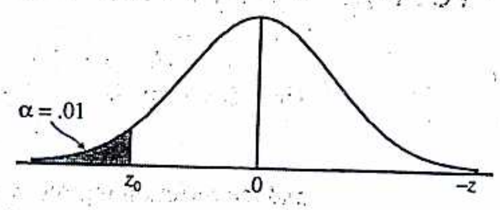

Figure 10.2

10.10 Let μ be the average hardness index. We are to test $H_0\colon\ \mu \geq 64$ vs. $H_a\colon\ \mu < 64$. The test statistic is

$$z = \frac{\overline{y}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{62-64}{\frac{8}{\sqrt{50}}} = -1.77$$

The rejection region is RR: Reject H_0 if $z < -z_{.01} = -2.326$.
Conclusion: Do not reject H_0 at $\alpha = .01$. There is insufficient evidence to reject the manufacturer's claim.

10.12 a. It might be reasonable to expect that the boys would have a greater interest in sports. Based on this we might choose $H_a\colon\ \mu_1 - \mu_2 > 0$, where μ_1 = mean LIC score for sports for boys and μ_2 = mean LIC score for sports for girls.

b. This corresponds to a one-tailed test.

c. The hypothesis of interest is

$$H_0\colon\ \mu_1 - \mu_2 = 0 \qquad \text{vs.} \qquad H_a\colon\ \mu_1 - \mu_2 > 0$$

and the test statistic is

$$z = \frac{(\overline{y}_1-\overline{y}_2)-0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{13.65-9.88}{\sqrt{\frac{(4.82)^2}{252}+\frac{(4.41)^2}{307}}} = 9.56$$

The rejection region, with $\alpha = .01$, is $z > 2.33$ and H_0 is rejected. There is evidence to indicate that the mean for men is greater than that for women.

6

10.14 a. If we define p as the proportion of college students aged 30 years or more, then we test

$$H_0\colon\ p = .25 \qquad \text{vs.} \qquad H_a\colon\ p \neq .25$$

The test statistic is

$$z = \frac{\widehat{p}-p_0}{\sqrt{\frac{p_0q_0}{n}}} = \frac{\frac{98}{300}-.25}{\sqrt{\frac{(.25)(.75)}{300}}} = 3.07$$

and the rejection region, with $\alpha = .05$ is $|z| > 1.96$. H_0 is rejected and we conclude that the 25% figure is not accurate.

b. Yes, the results do give evidence that the columnist's claim is too low.

10.16 Let p be the proportion of Americans with brown eyes. Then the hypothesis of interest is

$$H_0\colon\ p = .45 \qquad \text{vs.} \qquad H_a\colon\ p \neq .45$$

With $\widehat{p} = \frac{32}{80} = .40$ and $n = 80$, the test statistic is

$$z = \frac{\widehat{p}-p_0}{\sqrt{\frac{p_0q_0}{n}}} = \frac{.4-.45}{\sqrt{\frac{(.45)(.55)}{80}}} = -.90.$$

The rejection region is two-tailed with $\alpha = .01$, or $|z| > 2.58$ and H_0 is not rejected. There is insufficient evidence to indicate that the proportion of brown-eyed people in the region where the study was performed differs from the value reported in the *Washington Post.*

10.18 Throughout let p_1 be the relevant proportion in 1986 and p_2 be the appropriate proportion in 1991.

a. The hypothesis of interest is

$$H_0\colon\ p_1 - p_2 = 0 \qquad \text{vs.} \qquad H_a\colon\ p_1 - p_2 \neq 0$$

Calculate

$$\widehat{p}_1 = .45,\ \widehat{p}_2 = .34, \text{ and } \widehat{p} = \frac{n_1\widehat{p}_1+n_2\widehat{p}_2}{n_1+n_2} = \frac{450+340}{1000+1000} = .395$$

The test statistic is then

$$z = \frac{\widehat{p}_1-\widehat{p}_2}{\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{.45-.34}{\sqrt{(.395)(.605)\left(\frac{1}{1000}+\frac{1}{1000}\right)}} = 5.03.$$

The rejection region with $\alpha = .05$ is $z > 1.96$ and H_0 is rejected. There is evidence of a difference in the proportion of users in 1986 and 1991.

b. The hypothesis of interest is

$$H_0\colon\ p_1 - p_2 = 0 \qquad \text{vs.} \qquad H_a\colon\ p_1 - p_2 < 0$$

Calculate

$$\widehat{p} = .14,\ \widehat{p}_2 = .26, \text{ and } \widehat{p} = \frac{n_1\widehat{p}_1+n_2\widehat{p}_2}{n_1+n_2} = \frac{140+260}{1000+1000} = .2$$

The test statistic is then

$$z = \frac{\widehat{p}_1-\widehat{p}_2}{\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{.14-.26}{\sqrt{(.2)(.8)\left(\frac{1}{1000}+\frac{1}{1000}\right)}} = -6.71.$$

The rejection region with $\alpha = .05$ is $|z| < -1.645$ and H_0 is rejected. There is evidence to indicate that ibuprofen has significantly increased its market share from 1986 to 1991.

c. Yes. The survey is based on the same samples of 1000 people. Hence, if a person has decreased his use of aspirin, he may have begun using ibuprofen. The tests are not independent.