Jo Hardin
Math 152 – Fall 2008
HW # 11

8.8: 1, 2, 4
8.10: 18
9.1: 4, 5, 8, 9

1. Let's say we're watching taxis go by, and we're trying to decide if there are at least 4 per hour.

$$\begin{aligned}&X_i \sim \text{Poisson}(\theta)\\ &H_0 : \theta \leq 4\\ &H_1 : \theta > 4\end{aligned}$$

   Find the LRT for these hypotheses, assuming a sample of size n.

Solution: Note, MLE of $\theta = \hat{\theta} = \overline{X}$

$$\begin{aligned}
f(x_i|\theta) &= \frac{\theta^{x_i} e^{-\theta}}{x_i!}\\
f(\underline{x}|\theta) &= \frac{\theta^{\sum x_i} e^{-n\theta}}{\prod x_i!}\\
L(\hat{\theta}) &= \frac{\overline{x}^{\sum x_i} e^{-n\overline{x}}}{\prod x_i!}\\
L(\theta_0) &= \frac{4^{\sum x_i} e^{-n4}}{\prod x_i!}\\
\Lambda &= \frac{4^{\sum x_i} e^{-n4}}{\overline{x}^{\sum x_i} e^{-n\overline{x}}}\\
&= \left(\frac{4e}{\overline{x}}\right)^{\sum x_i} e^{-n4}
\end{aligned}$$

Our critical region: C $= \{(X_1, X_2, \ldots, X_n) | \left(\frac{4e}{\overline{x}}\right)^{\sum x_i} \leq k\}$ is a LRT.

If we find k such that $P(\underline{X} \in C | H_o \text{ true}) = 0.05$, then the critical region is a LRT of size α. How do we find k?

Let $g(a) = \left(\frac{4e}{a}\right)^{na} \Rightarrow g'(a) < 0 \Rightarrow$ as $a \to \infty$, $\left(\frac{4e}{a}\right)^{na}$ decreases.

So, our critical region becomes: C $= \{(X_1, X_2, \ldots, X_n) | \sum X_i \geq k'\}$ is a LRT.

To find k' we use $\sum X_i \sim$ Poisson or the CLT.

8.8

1. In this exercise, $\xi_0 = 0.9$, $\xi_1 = 0.1$, $w_0 = 1000$, and $w_1 = 18{,}000$. Also,

$$f_0(x) = \frac{1}{(2\pi)^{1/2}} \exp\left[-\frac{1}{2}(x-50)^2\right]$$

and

$$f_1(x) = \frac{1}{(2\pi)^{1/2}} \exp\left[-\frac{1}{2}(x-52)^2\right].$$

By the results of this section, it should be decided that the process is out of control if

$$\frac{f_1(x)}{f_0(x)} > \frac{\xi_0 w_0}{\xi_1 w_1} = \frac{1}{2}.$$

This inequality can be reduced to the inequality $2x - 102 > -\log\ 2$ or, equivalently, $x > 50.653$.

2. In this exercise, $\xi_0 = 2/3$, $\xi_1 = 1/3$, $w_0 = 1$, and $w_1 = 4$. Therefore, by the results of this section, it should be decided that f_0 is correct if

$$\frac{f_1(x)}{f_0(x)} < \frac{\xi_0 w_0}{\xi_1 w_1} = \frac{1}{2}.$$

Since $f_1(x)/f_0(x) = 4x^3$, it should be decided that f_0 is correct if $4x^3 < 1/2$ or, equivalently, if $x < 1/2$.

4. In this exercise, $\xi_0 = 1/4$, $\xi_1 = 3/4$, and $w_0 = w_1 = 1$. Let $x_1, \ldots, x_n$ denote the observed values in the sample, and let $y = \sum_{i=1}^{n} x_i$. Then

$$f_0(\boldsymbol{X}) = (0.3)^y (0.7)^{n-y}$$

and

$$f_1(\boldsymbol{X}) = (0.4)^y (0.6)^{n-y}.$$

By the results of this section, H_0 should be rejected if

$$\frac{f_1(\boldsymbol{X})}{f_0(\boldsymbol{X})} > \frac{\xi_0 w_0}{\xi_1 w_1} = \frac{1}{3}.$$

But

$$\frac{f_1(\boldsymbol{X})}{f_0(\boldsymbol{X})} = \left(\frac{4}{3} \cdot \frac{7}{6}\right)^y \left(\frac{6}{7}\right)^n > \frac{1}{3}$$

if and only if

$$y \log \frac{14}{9} + n \log \frac{6}{7} > \log \frac{1}{3}$$

or, equivalently, if and only if

$$\bar{x}_n = \frac{y}{n} > \frac{\log \frac{7}{6} + \frac{1}{n} \log \frac{1}{3}}{\log \frac{14}{9}}.$$

8.10

18. If U is defined as in Eq. (7.6.9), then the prior distribution of U is a t distribution with $2\alpha_0 = 2$ degrees of freedom. Since the t distribution is symmetric with respect to the origin, it follows that under the prior distribution, $\Pr(H_0) = \Pr(\mu \le 3) = \Pr(U \le 0) = 1/2$. It follows from (7.6.1) and (7.6.2) that under the posterior distribution,

$$\begin{aligned}
\mu_1 &= \frac{3 + (17)(3.2)}{1+17} = 3.189\,, \quad \lambda_1 = 18\,,\\
\alpha_1 &= 1 + \frac{17}{2} = 9.5\,,\\
\beta_1 &= 1 + \frac{1}{2}(17) + \frac{(17)(.04)}{2(18)} = 9.519.
\end{aligned}$$

If we now define Y to be the random variable in Eq. (7.6.12) then $Y = (4.24)(\mu - 3.19)$ and Y has a t distribution with $2\alpha_1 = 19$ degrees of freedom. Thus, under the posterior distribution,

$$\Pr(H_0) = \Pr(\mu \le 3) = \Pr[Y \le (4.24)(3 - 3.19)] = \Pr(Y \le -.81) = \Pr(Y \ge .81).$$

It is found from the table of the t distribution with 19 degrees of freedom that this probability is approximately 0.21.

9.1

4. We obtain the following table:

	AA	Aa	aa
N_i	10	10	4
np_i^0	6	12	6

It is found from Eq. (9.1.2) that $Q = 11/3$. If Q has a χ^2 distribution with 2 degrees of freedom, then the value of $\Pr(Q \ge 11/3)$ is between 0.1 and 0.2.

5. (a) The number of successes is $n\overline{X}_n$ and the number of failures is $n(1 - \overline{X}_n)$. Therefore,

$$\begin{aligned}
Q &= \frac{(n\overline{X}_n - np_0)^2}{np_0} + \frac{[n(1-\overline{X}_n) - n(1-p_0)]^2}{n(1-p_0)}\\
&= n(\overline{X}_n - p_0)^2 \left(\frac{1}{p_0} + \frac{1}{1-p_0}\right)\\
&= \frac{n(\overline{X}_n - p_0)^2}{p_0(1-p_0)}
\end{aligned}$$

(b) If $p = p_0$, then $E(\overline{X}_n) = p_0$ and $\operatorname{Var}(\overline{X}_n) = p_0(1-p_0)/n$. Therefore, by the central limit theorem, the d.f. of

$$Z = \frac{\overline{X}_n - p_0}{[p_0(1-p_0)/n]^{1/2}}$$

converges to the d.f. of the standard normal distribution. Since $Q = Z^2$, the d.f. of Q will converge to the d.f. of the χ^2 distribution with 1 degree of freedom.

8. If Z denotes a random variable having a standard normal distribution and X denotes the height of a man selected at random from the city, then

$$\begin{aligned}
\Pr(X < 66) &= \Pr(Z < -2) = 0.0227,\\
\Pr(66 < X < 67.5) &= \Pr(-2 < Z < -0.5) = 0.2858,\\
\Pr(67.5 < X < 68.5) &= \Pr(-0.5 < Z < 0.5) = 0.3830,\\
\Pr(68.5 < X < 70) &= \Pr(0.5 < Z < 2) = 0.2858,\\
\Pr(X > 70) &= \Pr(Z > 2) = 0.0227.
\end{aligned}$$

Therefore, we obtain the following table:

	N_i	np_i^0
$x < 66$	18	11.35
$66 < x < 67.5$	177	142.9
$67.5 < x < 68.5$	198	191.5
$68.5 < x < 70$	102	142.9
$x > 70$	5	11.35

It is found from Eq. (9.1.2) that $Q = 27.5$. If Q has a χ^2 distribution with 4 degrees of freedom, then $\Pr(Q \geq 27.5)$ is much less than 0.005.

9. (a) The five intervals, each of which has probability 0.2, are as follows:

$$(-\infty, -0.842),\ (-0.842, -0.253),\ (-0.253, 0.253),\ (0.253, 0.842),\ (0.842, \infty).$$

We obtain the following table:

	N_i	np_i^0
$-\infty < x < -0.842$	15	10
$-0.842 < x < -0.253$	10	10
$-0.253 < x < 0.253$	7	10
$0.253 < x < 0.842$	12	10
$0.842 < x < \infty$	6	10

The calculated value of Q is 5.4. If Q has a χ^2 distribution with 4 degrees of freedom, then $\Pr(Q \geq 5.4) = 0.25$.

(b) The ten intervals, each of which has probability 0.1, are as given in the following table:

	N_i	np_i^0
$-\infty < x < -1.282$	8	5
$-1.282 < x < -0.842$	7	5
$-0.842 < x < -0.524$	3	5
$-0.524 < x < -0.253$	7	5
$-0.253 < x < 0$	5	5
$0 < x < 0.253$	2	5
$0.253 < x < 0.524$	5	5
$0.524 < x < 0.842$	7	5
$0.842 < x < 1.282$	2	5
$1.282 < x < \infty$	4	5

The calculated value of Q is 8.8. If Q has a χ^2 distribution with 9 degrees of freedom, then the value of $\Pr(Q \geq 8.8)$ is between 0.4 and 0.5.