Assignment 4 | **LE, Nhat Hung**
MATH 324 - Statistics | McGill ID: 260793376
Prof. Masoud Asgharian | Date: March 28, 2019
Winter 2019 | Due date: March 29, 2019

10.95 p. 547
Suppose that we have a random sample of four observations from the density function

$$f(y|\theta) = \begin{cases} \left(\frac{1}{2\theta^3}\right) y^2 e^{-y/\theta}, & y > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

a. Find the rejection region for the most powerful test of $H_0 : \theta = \theta_0$ against $H_a : \theta = \theta_a$, assuming that $\theta_a > \theta_0$. [Hint: Make use of the χ^2 distribution.]

$$y_i \sim^{\text{iid}} \Gamma(\text{shape parameter } \alpha = 3, \text{scale parameter } \beta = \theta)$$

Because both hypotheses are simple, from the Neyman-Pearson Lemma, the most powerful test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \prod_i^4 \frac{f(y_i|\theta_0)}{f(y_i|\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^{12} e^{\left(\frac{1}{\theta_a} - \frac{1}{\theta_a}\right)\sum_i^4 y_i} < k$$

$$\Rightarrow \sum_i^4 y_i > \ln\left(k\left(\frac{\theta_0}{\theta_a}\right)^{12}\right)\left(\frac{1}{\theta_a} - \frac{1}{\theta_0}\right)^{-1}$$

Let

$$\ln\left(k\left(\frac{\theta_0}{\theta_a}\right)^{12}\right)\left(\frac{1}{\theta_a} - \frac{1}{\theta_0}\right)^{-1} = c$$

Under H_0,

$$\sum_i^4 y_i \sim \Gamma(\alpha = 4 \times 3 = 12, \beta = \theta_0)$$

$$\frac{2\sum_i^4 y_i}{\theta_0} \sim \chi^2(2\alpha = 24)$$

The rejection region of the best test can be rewritten as

$$\frac{2\sum_i^4 y_i}{\theta_0} > \frac{2c}{\theta_0} = c^*$$

where c^* can be chosen from the χ^2 distribution table.

b. Is the test given in part (a) uniformly most powerful for the alternative $\theta > \theta_0$?

Because the rejection region doesn't depend on a $\theta_a < \theta_0$, the test in part (a) is uniformly most powerful for the alternative $\theta > \theta_0$.

10.97 p. 547

Let $Y_1, Y_2, \ldots, Y_n$ be independent and identically distributed random variables with discrete probability function given by

	y		
	1	2	3
$p(y \mid \theta)$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

where $0 < \theta < 1$. Let N_i denote the number of observations equal to *i* for $i = 1, 2, 3$.

a. Derive the likelihood function $L(\theta)$ as a function of N_1, N_2 and N_3.

$$Y_i \sim^{\text{iid}} \text{Multinomial}(n = 3)$$

Therefore,

$$L(\theta) = \frac{n!}{N_1!N_2!N_3!}\theta^{2N_1}(2\theta(1-\theta))^{N_2}(1-\theta)^{2N_3}$$

b. Find the most powerful test for testing $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$, where $\theta_a > \theta_0$. Show that your test specifies that H_0 be rejected for certain values of $2N_1 + N_2$.

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_0}{\theta_a}\right)^{2N_1+N_2}\left(\frac{1-\theta_0}{1-\theta_a}\right)^{N_2+2N_3} < k$$

$$\Rightarrow \frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_0}{\theta_a}\right)^{2N_1+N_2}\left(\frac{1-\theta_0}{1-\theta_a}\right)^{2n-(2N_1+N_2)} < k$$

$$\Rightarrow (2N_1+N_2)\ln\frac{\theta_0}{\theta_a} + (2n-(2N_1+N_2))\ln\frac{1-\theta_0}{1-\theta_a} < \ln k$$

$$\Rightarrow 2N_1 + N_2 > \left(\ln k - 2n\ln\frac{1-\theta_0}{1-\theta_a}\right)\left(\ln\frac{\theta_0(1-\theta_a)}{\theta_a(1-\theta_0)}\right)^{-1} = c$$

The rejection region is therefore

$$\{2N_1 + N_2 > c\}$$

c. How do you determine the value of *k* so that the test has nominal level *α*? You need not do the actual computation. A clear description of how to determine *k* is adequate.

Find the distribution of $2N_1 + N_2$, find critical value *c* such that $P(2N_1 + N_2 > c) = \alpha$, solve c = *critical value* for *k*.

d. Is the test derived in parts (a)–(c) uniformly most powerful for testing $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$? Why or why not?

Because the RR doesn't depend on any specific θ_a, we have an UMP test.

10.106 p. 554

A survey of voter sentiment was conducted in four midcity political wards to compare the fraction of voters favoring candidate A. Random samples of 200 voters were polled in each of the four wards, with the results as shown in the accompanying table. The numbers of voters favoring A in the four samples can be regarded as four independent binomial random variables. Construct a likelihood ratio test of the hypothesis that the fractions of voters favoring candidate A are the same in all four wards. Use $\alpha = .05$.

	Ward				
Opinion	1	2	3	4	Total
Favor A	76	53	59	48	236
Do not favor A	124	147	141	152	564
Total	200	200	200	200	800

Let p_i be the probability of favoring candidate A in ward i, and y_i the number of people favoring candidate A in ward i.

$H_0 : p_1 = p_2 = p_3 = p_4 = p$

Let $\Theta = \{p_1, p_2, p_3, p_4\}$

$$L(\Theta) = \prod_i^4 \binom{200}{y_i} p_i^{y_i}(1-p_i)^{200-y_i}$$

Under H_0,

$$L(\Omega_0) = \left[\prod \binom{200}{y_i}\right] p^{\sum y_i}(1-p)^{800-\sum y_i}$$

$$\Rightarrow \ln L(\Omega_0) = \left[\sum \ln \binom{200}{y_i}\right] + \left[\sum y_i\right] \ln p + (800 - \sum y_i)\ln(1-p)$$

$$\Rightarrow \frac{\partial \ln L(\Omega_0)}{\partial p} = \frac{\sum y_i}{p} - \frac{800 - \sum y_i}{1-p}$$

$$\frac{\partial \ln L(\Omega_0)}{\partial p} = 0 \Rightarrow p = \frac{\sum y_i}{800} \Rightarrow \hat{\Omega}_0 = \frac{\sum y_i}{800}$$

For the unrestricted space,

$$\ln L(\Omega) = \sum_i \ln \binom{200}{y_i} + \left[\sum_i y_i\right] \ln p_i + \sum_i (200 - y_i)\ln(1-p_i)$$

$$\Rightarrow \frac{\partial \ln L(\Omega)}{\partial p_j} = \frac{y_j}{p_j} - \frac{200 - y_j}{1-p_j}$$

$$\Rightarrow \frac{\partial \ln L(\Omega)}{\partial p_j} = 0 \Rightarrow p_j = \frac{y_j}{200} \Rightarrow \hat{\Omega} = \frac{y_j}{p_j}$$

Then, the likelihood ratio is

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{\left(\frac{\sum y_i}{800}\right)^{\sum y_i}\left(1 - \frac{\sum y_i}{800}\right)^{800-\sum y_i}}{\prod_i^4 \left(\frac{y_j}{200}\right)^{y_i}\left(1 - \frac{y_i}{200}\right)^{200-y_i}}$$

We have a large sample size, with $r_0 = 3$ free parameters are fixed in Ω_0, and $r = 0$ in Ω. Therefore, according to Theorem 10.2 p. 551,

$$-2\ln\lambda \sim \chi^2(\nu = r_0 - r = 3)$$

and we reject H_0 if $-2\ln\lambda > \chi^2_{.05}$, with $\chi^2_{.05} = 7.81$.

From the data table above,

$$-2\ln\lambda = 10.54 \Rightarrow -2\ln\lambda > \chi^2_{.05}$$

Therefore, we reject H_0 and conclude that the fractions of voters favoring candidate are not the same in all 4 wards.

10.111 p. 555

Suppose that we are interested in testing the *simple* null hypothesis $H_0 : \theta = \theta_0$ versus the *simple* alternative hypothesis $H_a : \theta = \theta_a$. According to the Neyman–Pearson lemma, the test that maximizes the power at θ_a has a rejection region determined by

$$\frac{L(\theta_0)}{L(\theta_a)} < k.$$

In the context of a likelihood ratio test, if we are interested in the *simple* H_0 and H_a, as stated, then $\Omega_0 = \{\theta_0\}, \Omega_a = \{\theta_a\}$, and $\Omega = \{\theta_0, \theta_a\}$.

a. Show that the likelihood ratio λ is given by

$$\lambda = \frac{L(\theta_0)}{\max(L(\theta_0), L(\theta_a))} = \frac{1}{\max\left\{1, \frac{L(\theta_a)}{L(\theta_0)}\right\}}.$$

Under H_0, the MLE is θ_0, and under the alternative hypothesis, the MLE can be either θ_0 or θ_a. Therefore, $L(\widehat{\Omega}_0) = L(\theta_0)$ and $L(\widehat{\Omega}) = \max(L(\theta_0), L(\theta_a))$.

b. Argue that $\lambda < k$ if and only if, for some constant k',

$$\frac{L(\theta_0)}{L(\theta_a)} < k'.$$

$$\lambda = \frac{1}{\max\left\{1, \frac{L(\theta_a)}{L(\theta_0)}\right\}} = \min\left\{1, \frac{L(\theta_0)}{L(\theta_a)}\right\}$$

Therefore, $\lambda < k < 1$ iff $L(\theta_0)/L(\theta_a) < k$.

c. What do the results in parts (a) and (b) imply about likelihood ratio tests when both the null and alternative hypotheses are simple?

Using a likelihood ratio test will result in the same rejection region as when using the Neyman-Pearson Lemma.

11.14 p. 576

J. H. Matis and T. E. Wehrly report the following table of data on the proportion of green sunfish that survive a fixed level of thermal pollution for varying lengths of time.

Proportion of Survivors (y)	Scaled Time (x)
1.00	.10
.95	.15
.95	.20
.90	.25
.85	.30
.70	.35
.65	.40
.60	.45
.55	.50
.40	.55

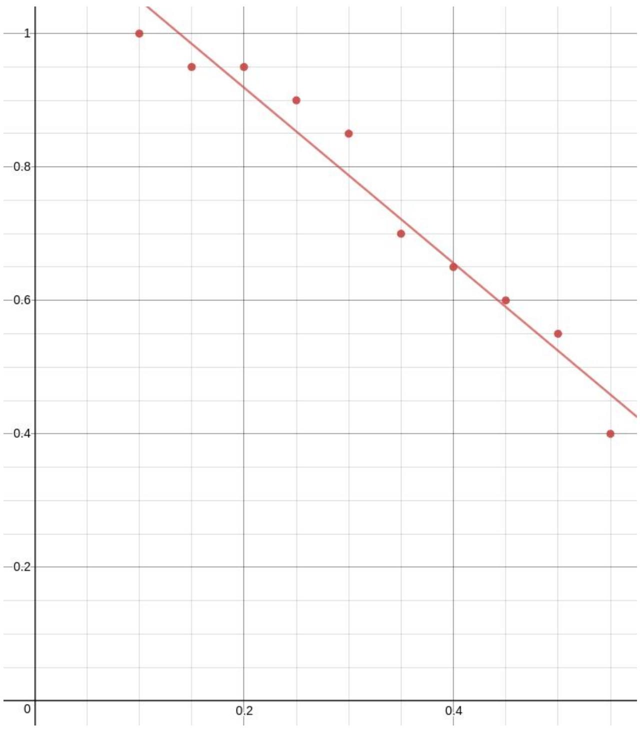
a. Fit the linear model $Y = \beta_0 + \beta_1 x + \varepsilon$. Give your interpretation.

$$\overline{x} = 0.325, \overline{y} = 0.775,$$
$$\beta_1 = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} \approx -1.315, \beta_0 = \overline{y} - \beta_1 \approx 1.182$$

Therefore, the least squares line is

$$y = 1.182 - 1.315x$$

b. Plot the points and graph the result of part (a). Does the line fit through the points?



The line fits pretty well.

11.20 p. 583

Suppose that $Y_1, Y_2, ..., Y_n$ are independent normal random variables with $E(Y_i) = \beta_0 + \beta_1 x_i$ and $V(Y_i) = \sigma^2$, for $i = 1, 2, ..., n$. Show that the maximum-likelihood estimators (MLEs) of β_0 and β_1 are the same as the least-squares estimators of Section 11.3.

$$L(\beta_0, \beta_1) = (\sigma\sqrt{2\pi})^{-n}\text{exp}\left\{-\frac{1}{2\sigma^2}\sum(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$
$$\Rightarrow \ln L(\beta_0, \beta_1) = -n\ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum(y_i - \beta_0 - \beta_1 x_i)^2$$
$$\text{maximize } \ln L(\beta_0, \beta_1) \Rightarrow \text{minimize } \sum(y_i - \beta_0 - \beta_1 x_i)^2$$

which is the least-squares condition. Therefore, the MLEs of β_0 and β_1 are the same as the least-squares estimators: $\widehat{\beta}_0$ and $\widehat{\beta}_1$.

11.22 p. 583
Under the assumptions of Exercise 11.20, find the MLE of σ^2 .

$$\ln L(\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum(y_i - \beta_0 - \beta_1 x_i)^2$$
$$\frac{\partial \ln L(\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(y_i - \beta_0 - \beta_1 x_i)^2$$
$$\frac{\partial \ln L(\sigma^2)}{\partial \sigma^2} = 0 \Rightarrow \sigma^2 = \frac{1}{n}\sum(y_i - \beta_0 - \beta_1 x_i)^2$$

Plugging in the MLEs of β_0 and β_1 , we have the MLE of σ^2 equal to
$$\hat{\sigma}^2 = \frac{1}{n}\sum(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = \text{SSE}/n$$

11.27 p. 587
Use the properties of the least-squares estimators given in Section 11.4 to complete the following.
a. Show that under the null hypothesis $H_0 : \beta_i = \beta_{i0}$
$$T = \frac{\hat{\beta}_i - \beta_{i0}}{S\sqrt{c_{ii}}}$$
Possesses a *t* distribution with $n-2$ df, where $i = 1, 2$.

$$Z = \frac{\hat{\beta}_i - \beta_{i0}}{\sigma\sqrt{c_{ii}}} \sim \text{Normal}(\mu = 0, \sigma^2 = 1) \text{ (standard)}$$
$$W = \frac{(n-2)S^2}{\sigma^2} \sim \chi^2(\nu = n-2)$$

Therefore, from definition 7.2 p. 360
$$\frac{Z}{\sqrt{W/\nu}} = \frac{Z}{\sqrt{W/(n-2)}} = \frac{\hat{\beta}_i - \beta_{i0}}{S\sqrt{c_{ii}}} = T \sim t(\nu = n-2)$$

b. Derive the confidence intervals for β_i given in this section.

The CI, using *T* as the pivotal value, is

$$\hat{\beta}_i \pm (t_{\alpha/2} S\sqrt{c_{ii}} + \beta_{i0})$$

11.30 p. 588

The octane number Y of refined petroleum is related to the temperature x of the refining process, but it is also related to the particle size of the catalyst. An experiment with a small-particle catalyst gave a fitted least-squares line of

$$\hat{y} = 9.360 + .155x,$$

with $n = 31$, $V(\widehat{\beta}_1) = (.0202)^2$, and $\text{SSE} = 2.04$. An independent experiment with a large-particle catalyst gave

$$\hat{y} = 4.265 + .190x,$$

with $n = 11$, $V(\widehat{\beta}_1) = (.0193)^2$, and $\text{SSE} = 1.86$.

a. Test the hypotheses that the slopes are significantly different from zero, with each test at the significance level of .05.

$H_0 : \beta_1 = 0$, $H_a : \beta_1 \neq 0$.

$$T_1 = \frac{.155}{.0202} \approx 7.67 \sim t(\nu = 31 - 2 = 29)$$

At 29 df, $t_{.025} = 2.045$. $T_1 > t_{.025}$, so we reject H_0 for the first slope.

$$T_2 = \frac{.190}{.0193} \approx 9.84 \sim t(\nu = 11 - 2 = 9)$$

At 9 df, $t_{.025} = 2.262$. $T_2 > t_{.025}$, so we also reject H_0 for the second slope.

11.40 p. 593

Refer to Exercise 11.14. Find a 90% confidence interval for the expected proportion of survivors at time period .30.

From 11.14, the least-squares line is

$$y = 1.182 - 1.315x$$

Therefore, at time period $x = .30$,

$$\hat{y} = 1.182 - (1.315)(.30) = .7875$$

The sample size is $n = 10$, so our critical value $t_{.05} = 1.86$ is based on $n - 2 = 8$ df.

Furthermore,

$$\overline{x} = .325$$
$$S_{xx} = .20625$$
$$S^2 = \frac{\text{SEE}}{n-2} = \frac{S_{yy} - \hat{\beta}_1 S_{xy}}{8} = \frac{.0155}{8}$$

Therefore, the CI for $E(Y)$ is

$$\hat{y} \pm t_{.05} S\sqrt{\frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}} = .7875 \pm 1.86\sqrt{\frac{.0155}{8}\left(\frac{1}{10} + \frac{(.3 - .325)^2}{.20625}\right)}$$
$$\approx (.761, .814)$$

11.47 p. 597

Refer to Exercise 11.14. Find a 95% prediction interval for the proportion of survivors at time $x = .60$.

When $x = .60$, $\widehat{y} = .393$. The critical value is $t_{.025} = 2.306$. Therefore, the 95% prediction interval (PI) for the proportion of survivors at time $x = .60$ is.

$$.393 \pm 2.306\sqrt{\frac{.0155}{8}\left(1+\frac{1}{10}+\frac{(.6-.325)^2}{.20625}\right)} \approx (.270, .516)$$

11.55 p. 603

Consider the simple linear regression model based on normal theory. If we are interested in testing $H_0 : \beta_1 = 0$ versus various alternatives, the statistic

$$T = \frac{\hat{\beta}_1 - 0}{S/\sqrt{S_{xx}}}$$

possesses a t distribution with $n-2$ df if the null hypothesis is true. Show that the equation for T can also be written as

$$T = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

$$T = \frac{\hat{\beta}_1}{S} = \frac{\hat{\beta}_1\sqrt{S_{xx}}}{\sqrt{\mathrm{SSE}/(n-2)}} = \frac{\hat{\beta}_1\sqrt{S_{xx}}\sqrt{n-2}}{\sqrt{S_{yy}-\hat{\beta}_1 S_{xy}}} = \frac{\hat{\beta}_1\sqrt{S_{xx}/S_{yy}}\sqrt{n-2}}{\sqrt{1-\frac{S_{xy}}{S_{xx}}\frac{S_{xy}}{S_{yy}}}}$$

$$= \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$