STAT420 Homework 5

James Zhang*

October 20, 2023

*Email: jzhang72@terpmail.umd.com

4.5.4. Let X have a binomial distribution with the number of trials $n = 10$ and with p either 1/4 or 1/2. The simple hypothesis $H_0 : p = \frac{1}{2}$ is rejected, and the alternative simple hypothesis $H_1 : p = \frac{1}{4}$ is accepted, if the observed value of X_1, a random sample of size 1, is less than or equal to 3. Find the significance level and the power of the test.

Recall that the α is sometimes called the significance level, where

$$\alpha = P(X_1 \le 3 | H_0) = \sum_{i=0}^{3} \binom{10}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{10-i} \approx 0.1719$$

Now the power of the test is

$$1 - \beta = P(X_1 \le 3 | H_1) = \sum_{i=0}^{3} \binom{10}{i} (0.25)^i (0.75)^{10-i} \approx 0.7759$$

4.5.8. Let us say the life of a tire in miles, say X, is normally distributed with mean θ and standard deviation 5000. Past experience indicates that $\theta = 30{,}000$. The manufacturer claims that the tires made by a new process have mean $\theta > 30{,}000$. It is possible that $\theta = 35{,}000$. Check his claim by testing $H_0 : \theta = 30{,}000$ against $H_1 : \theta > 30{,}000$. We observe n independent values of X, say $x_1, \ldots, x_n$, and we reject H_0 (thus accept H_1) if and only if $\overline{x} \ge c$. Determine n and c so that the power function $\gamma(\theta)$ of the test has the values $\gamma(30{,}000) = 0.01$ and $\gamma(35{,}000) = 0.98$.

We will create a system of equations using conditions of the power function and properties of the standard normal distribution. We have

$$\frac{c - 30000}{5000/\sqrt{n}} = z_{0.01} = P(Z > x) \implies x \approx 2.326$$

$$\frac{c - 35000}{5000/\sqrt{n}} = z_{0.98} \approx -2.054$$

Solving the system

$$\frac{c - 30000}{5000/\sqrt{n}} - \frac{c - 35000}{5000/\sqrt{n}} = z_{0.01} - z_{0.98} \implies \sqrt{n} \approx 4.38 \implies n \approx 19.194 \implies n = 20$$

Plugging this result back into our first equation we obtain

$$\frac{4.38(c - 30000)}{5000} 2.326 \implies c \approx 32655.25$$

4.7.4. A die was cast $n = 120$ independent times and the following data resulted:

Spots Up	1	2	3	4	5	6
Frequency	b	20	20	20	20	$40 - b$

If we use a chi-square test, for what values of b would the hypothesis that the die is unbiased be rejected at the 0.025 significance level?

We seek r such that $P(\chi^2 \leq r) = 1 - 0.025 = 0.975$ given that there are 5 degrees of freedom. Looking at Appendix D in the textbook, we get $r = 12.833$. Thus, if the result of

$$\text{(Observed - Expected)}^2/\text{Expected} > 12.833$$

then we can reject the null hypothesis. Therefore, we seek b such that

$$\frac{(b-20)^2}{20} + 0 + 0 + 0 + 0 + \frac{(40-b-20)^2}{20} > 12.833$$

$$\frac{(b-20)^2}{10} > 12.833 \implies b > \sqrt{128.33} + 20 \text{ and } b < -\sqrt{128.33} + 20$$

Thus $b > 31.328$ and $b < 8.672$ would cause the data to be rejected at the 0.025 significance level.

4.7.6. Two different teaching procedures were used on two different groups of students. Each group contained 100 students of about the same ability. At the end of the term, an evaluating team assigned a letter grade to each student. The results were tabulated as follows.

	Grade					
Group	A	B	C	D	F	Total
I	15	25	32	17	11	100
II	9	18	29	28	16	100

If we consider these data to be independent observations from two respective multinomial distributions with $k = 5$, test at the 5% significance level the hypothesis that the two distributions are the same (and hence the two teaching procedures are equally effective). For computation in R, use

```
r1=c(15,25,32,17,11);r2=c(9,18,29,28,16);mat=rbind(r1,r2)
chisq.test(mat)
```

A 5% significance level and $k = 5$ means we seek r such that $P(\chi^2_4 \leq r) = 0.95 \implies r = 9.488$. Thus, if our $Q_{k-1} > 9.488$, then we reject the null hypothesis and conclude that the distributions are not homogeneous. Now let us compute Q_{k-1}

$$Q_{k-1} = 2(\frac{(15-12)^2}{12} + \frac{(25-21.5)^2}{21.5} + \frac{(32-30.5)^2}{30.5} + \frac{(17-22.5)^2}{22.5} + \frac{(11-13.5)^2}{13.5})$$
$$\approx 6.4109$$

and so we fail to reject the null hypothesis and conclude that the distributions are the same.

4.7.8. Let the result of a random experiment be classified as one of the mutually exclusive and exhaustive ways A_1, A_2, A_3 and also as one of the mutually exhaustive ways B_1, B_2, B_3, B_4. Say that 180 independent trials of the experiment result in the following frequencies:

	B_1	B_2	B_3	B_4
A_1	$15-3k$	$15-k$	$15+k$	$15+3k$
A_2	15	15	15	15
A_3	$15+3k$	$15+k$	$15-k$	$15-3k$

where k is one of the integers $0, 1, 2, 3, 4, 5$. What is the smallest value of k that leads to the rejection of the independence of the A attribute and the B attribute at the $\alpha = 0.05$ significance level?

Given a 0.05 significance level and $r = 3, c = 4 \implies (r-1)(c-1) = 6$, we seek x such that $P(\chi^2_6 \leq x) = 0.95 \implies x = 12.592$, so if our test statistic is greater than 12.592, we reject the null hypothesis. We will proceed by calculating the test statistic where the expected count is row sum times column sum divided by total. Note that all row sums and column sums are 60 and 45, respectively, so all expected counts equal

$$\frac{60(45)}{180} = 15$$

Now we calculate

$$4\frac{(15-3k-15)^2}{15} + 4\frac{(15-k-15)^2}{15} = \frac{40k^2}{15}$$

Note that there are 2 occurrences of $15+3k$ and $15-3k$, but the 15 cancels off when subtracting the expected count, and $(-3k)^2 = (3k)^2$. So we have 4 occurrences of "3k terms", 4 occurrences of "k terms", and 4 occurrences of "0 terms". Continuing on, we seek the smallest $k \in \{0, 1, 2, 3, 4, 5\}$ such that

$$\frac{40k^2}{15} > 12.592$$

$$k^2 > \frac{12.592(15)}{40} \implies k > 2.17 \implies$$

the smallest k is 3.

6.3.6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu_0, \sigma^2 = \theta)$ distribution, where $0 < \theta < \infty$ and μ_0 is known. Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ can be based upon the statistic $W = \sum_{i=1}^n (X_i - \mu_0)^2/\theta_0$. Determine the null distribution of W and give, explicitly, the rejection rule for a level α test.

Note the likelihood function is

$$L(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} e^{-(x_i-\mu_0)^2/2\theta} \implies L(\theta_0) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_0}} e^{-W/2} = (\frac{1}{\sqrt{2\pi\theta_0}})^n e^{-W/2}$$

We seek the maximum likelihood estimator of θ, $\hat{\theta}$. The log likelihood function is

$$l(\theta) = \sum_{i=1}^{n} \log f(X_i; \theta) = \sum_{i=1}^{n} \log \frac{1}{\sqrt{2\pi\theta}} e^{-(x_i-\mu_0)^2/2\theta}$$

$$= \sum_{i=1}^{n} -\frac{1}{2}\log(2\pi\theta) - (x_i - \mu_0)^2/2\theta = -\frac{n}{2}\log(2\pi\theta) - \sum_{i=1}^{n} \frac{(x_i - \mu_0)^2}{2\theta}$$

Taking the first derivative

$$l'(\theta) = -\frac{n}{2\theta} + \sum_{i=1}^{n} \frac{(x_i - \mu_0)^2}{2\theta^2}$$

Setting the first derivative equal to 0, we obtain

$$\sum_{i=1}^{n} \frac{(x_i - \mu_0)^2}{2\theta^2} = \frac{n}{2\theta}$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_0)^2$$

Thus, plugging this result into the likelihood function

$$L(\hat{\theta}) = (\frac{1}{\sqrt{2\pi \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_0)^2}})^n e^{-\sum_{i=1}^{n} (x_i-\mu_0)^2/2(\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2)}$$

$$= (\frac{1}{\sqrt{2\pi \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_0)^2}})^n e^{-\frac{n}{2}}$$

The likelihood ratio test statistic is

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta})} = \frac{(\frac{1}{\sqrt{2\pi\theta_0}})^n e^{-W/2}}{(\frac{1}{\sqrt{2\pi\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2}})^n e^{-\frac{n}{2}}} = (\frac{W}{n})^{n/2} e^{-\frac{1}{2}(W+n)}$$

and so this likelihood ratio test clearly depends on W, as desired. It is clear by the definition of chi squared distribution that $W \sim \chi^2$. By a theorem, the null distribution of W is chi-squared with 1 degree of freedom. Now note that $\Lambda \leq c \implies W \leq c_1$ or $W \geq c_2$ for constants c_1, c_2. We take $c_1 = \chi^2\alpha/2(n)$ and $c_2 = \chi^2_{1-\alpha/2}(n)$ to get a level α test.

6.3.9. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean $\theta > 0$.

(a) Show that the likelihood ratio test of $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$ is based upon the statistic $Y = \sum_{i=1}^{n} X_i$. Obtain the null distribution of Y.

(b) For $\theta_0 = 2$ and $n = 5$, find the significance level of the test that rejects H_0 if $Y \leq 4$ or $Y \geq 17$.

a. The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f(\theta) = \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta} = \frac{\theta^{x_i}}{x_i!} e^{-n\theta} \implies L(\theta_0) = \frac{\theta_0^{\sum x_i}}{x_i!} e^{-n\theta_0}$$

By a theorem, the MLE of a sample from a Poisson distribution is the sample mean so $\hat{\theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and so

$$L(\hat{\theta}) = \frac{(\frac{1}{n}\sum_{i=1}^{n} X_i)^{x_i}}{x_i!} e^{-n(\frac{1}{n}\sum_{i=1}^{n} X_i)} = \frac{(Y/n)^{\sum x_i}}{x_i!} e^{-Y}$$

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta})} = (\theta_0 n/Y)^Y \exp(Y - n\theta_0)$$

and so it becomes clear that this test is dependent on the statistic Y. By a theorem, the sum of Poisson random variables results in a Poisson random variable with the sum of the parameters, so the null distribution of Y is Poisson$(n\theta_0)$, given that each Poisson distribution in the sample has parameter θ_0.

b. We seek to find the significance level α for $Y \sim$ Poisson$(n\theta_0) =$ Poisson(10).

$$\alpha = P(Y \leq 4) + P(Y \geq 17) = 0.02925 + 0.02704 = 0.05629$$

6.3.13. Let $X_1, X_2, \ldots, X_n$ be a random sample from the beta distribution with $\alpha = \beta = \theta$ and $\Omega = \{\theta : \theta = 1, 2\}$. Show that the likelihood ratio test statistic Λ for testing $H_0 : \theta = 1$ versus $H_1 : \theta = 2$ is a function of the statistic $W = \sum_{i=1}^{n} \log X_i + \sum_{i=1}^{n} \log(1 - X_i)$.

Recall that the pdf of a beta distribution is

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx} \implies f(x; \theta, \theta) = \frac{x^{\theta-1}(1-x)^{\theta-1}}{\int_0^1 x^{\theta-1}(1-x)^{\theta-1} dx}$$

The likelihood function

$$L(\theta) = \prod_{i=1}^{n} f(x; \theta; \theta) = \prod_{i=1}^{n} \frac{x_i^{\theta-1}(1-x_i)^{\theta-1}}{\int_0^1 x_i^{\theta-1}(1-x_i)^{\theta-1} dx} \implies L(1) = \prod_{i=1}^{n} \frac{1}{x_i}$$

$$L(2) = \prod_{i=1}^{n} \frac{x_i(1-x_i)}{\int_0^1 x_i - x_i^2 dx} = \prod_{i=1}^{n} \frac{x_i(1-x_i)}{\frac{1}{2}x_i^2 - \frac{1}{3}x_i^3} = \prod_{i=1}^{n} \frac{(1-x_i)}{\frac{1}{2}x_i - \frac{1}{3}x^2}$$

Thus, our test statistic is

$$\Lambda = \frac{L(1)}{L(2)} = \prod_{i=1}^{n} \frac{\frac{1}{2}x_i - \frac{1}{3}x^2}{x_i(1-x_i)} = \sum_{i=1}^{n} \log \frac{\frac{1}{2}x_i - \frac{1}{3}x^2}{x_i(1-x_i)}$$

$$= \sum_{i=1}^{n} \log(\frac{1}{2}x_i - \frac{1}{3}x^2) - \sum_{i=1}^{n} \log(x_i) - \sum_{i=1}^{n} \log(1-x_i)$$

$$\Lambda = \sum_{i=1}^{n} \log(\frac{1}{2}x_i - \frac{1}{3}x_i^2) - W$$

and so we conclude that the likelihood ration test statistic Λ is a function of the statistic W, as desired.

6.3.18. Let $X_1, X_2, \ldots, X_n$ be a random sample from a $\Gamma(\alpha, \beta)$ distribution where α is known and $\beta > 0$. Determine the likelihood ratio test for $H_0 : \beta = \beta_0$ against $H_1 : \beta \neq \beta_0$.

Recall that the pdf of a gamma distribution is

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha \int_0^\infty x^{\alpha-1}e^{-x}\ dx}$$

where the denominator is the function $\Gamma(\alpha)$. The likelihood function $L(\theta)$ is

$$L(\beta) = \prod_{i=1}^{n} f(x; \alpha, \beta) \implies L(\beta_0) = \prod_{i=1}^{n} \frac{x_i^{\alpha-1}e^{-x_i/\beta_0}}{\beta_0^\alpha \int_0^\infty x_i^{\alpha-1}e^{-x_i}\ dx}$$

Now we seek the maximum likelihood estimator of β, $\hat{\beta}$ so we take the log likelihood

$$l(\beta) = \sum_{i=1}^{n} \log f(x; \alpha, \beta) = \sum_{i=1}^{n} \log \frac{x_i^{\alpha-1}e^{-x_i/\beta}}{\beta^\alpha \int_0^\infty x_i^{\alpha-1}e^{-x_i}\ dx}$$

$$l(\beta) = \sum_{i=1}^{n} \log(x_i^{\alpha-1}) + \sum_{i=1}^{n} -x_i/\beta - \sum_{i=1}^{n} \log(\beta^\alpha) - \sum_{i=1}^{n} \log(\int_0^\infty x_i^{\alpha-1}e^{-x_i}dx)$$

Taking the first derivative with respect to β we get

$$l'(\beta) = \sum_{i=1}^{n} \frac{x_i}{\beta^2} - \frac{\alpha\beta^{\alpha-1}}{\beta^\alpha}n = \frac{1}{\beta^2}\sum_{i=1}^{n} x_i - \frac{\alpha n}{\beta}$$

by using power rule and chain rule. Setting the first derivative to 0 and solving for β

$$\frac{1}{\beta^2}\sum_{i=1}^{n} x_i = \frac{\alpha n}{\beta} \implies \hat{\beta} = \frac{1}{\alpha n}\sum_{i=1}^{n} x_i = \frac{\bar{X}}{\alpha}$$

Plugging this result back into our likelihood function we get

$$L(\hat{\beta}) = \prod_{i=1}^{n} \frac{x_i^{\alpha-1}e^{-x_i/(\frac{\bar{X}}{\alpha})}}{(\frac{\bar{X}}{\alpha})^\alpha \int_0^\infty x_i^{\alpha-1}e^{-x_i}\ dx}$$

$$\Lambda = \frac{L(\beta_0)}{L(\hat{\beta})} = \frac{\prod_{i=1}^{n} \frac{x_i^{\alpha-1}e^{-x_i/\beta_0}}{\beta_0^\alpha \int_0^\infty x_i^{\alpha-1}e^{-x_i}\ dx}}{\prod_{i=1}^{n} \frac{x_i^{\alpha-1}e^{-x_i/(\frac{\bar{X}}{\alpha})}}{(\frac{\bar{X}}{\alpha})^\alpha \int_0^\infty x_i^{\alpha-1}e^{-x_i}\ dx}} = \prod_{i=1}^{n} \frac{e^{-x_i/\beta_0}(\bar{X}/\alpha)^\alpha}{\beta_0^\alpha e^{-x_i/(\frac{\bar{X}}{\alpha})}}$$

$$\Lambda = \prod_{i=1}^{n} (\frac{\bar{X}}{\alpha\beta_0})^{\alpha} e^{-x_i/\beta_0 + x_i/(\frac{\bar{X}}{\alpha})} = \prod_{i=1}^{n} (\frac{\bar{X}}{\alpha\beta_0})^{\alpha} \exp(x_i(\frac{\alpha}{\bar{X}} - \frac{1}{\beta_0}))$$
$$= (\frac{\bar{X}}{\alpha\beta_0})^{\alpha n} \sum_{i=1}^{n} x_i(\frac{\alpha}{\bar{X}} - \frac{1}{\beta_0}) = (\frac{\bar{X}}{\alpha\beta_0})^{\alpha n}(n\bar{X})(\frac{\alpha}{\bar{X}} - \frac{1}{\beta_0})$$

6.5.1. On page 80 of their test, Hollander and Wolfe (1999) present measurements of the ratio of the earth's mass to that of its moon that were made by 7 different spacecraft (5 of the Mariner type and 2 of the Pioneer type). These measurements are presented below (also in the file `earthmoon.rda`). Based on earlier Ranger voyages, scientists had set this ratio at 81.3035. Assuming a normal distribution, test the hypotheses $H_0 : \mu = 81.3035$ versus $H_1 : \mu \neq 81.3035$, where μ is the true mean ratio of these later voyages. Using the p-value, conclude in terms of the problem at the nominal α-level of 0.05.

Earth to Moon Mass Ratios						
81.3001	81.3015	81.3006	81.3011	81.2997	81.3005	81.3021

Note that we that the sample is from a normal distribution, we want to test the mean, and the variance is unknown, so we will proceed with a $t-$test.

$$\left| \frac{\sqrt{n}(\bar{X} - \mu_0)}{S} \right| > c$$

then we reject the null hypothesis. Note that $c = t_{\alpha/2,n-1} = t_{0.025,6} = 2.4469$. We compute the following values such that

$$\bar{X} = 81.3008, \sqrt{n} = \sqrt{7}, \mu_0 = 81.3035, \text{ and}$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} = 0.000827$$

and so

$$\left| \frac{\sqrt{7}(81.3008 - 81.3035)}{0.000827} \right| = 8.64 > 2.4469$$

so we reject the null hypothesis.

6.5.4. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution $N(\theta_1, \theta_2)$. Show that the likelihood ratio principle for testing $H_0 : \theta_2 = \theta_2'$ specified, and θ_1 unspecified against $H_1 : \theta_2 \neq \theta_2'$, θ_1 unspecified, leads to a test that rejects when $\sum_1^n (x_i - \overline{x})^2 \leq c_1$ or $\sum_1^n (x_i - \overline{x})^2 \geq c_2$, where $c_1 < c_2$ are selected appropriately.

By a theorem using MME, the maximum likelihood estimators of the mean and variance of a normal distribution are the sample mean and sample variance, respectively. Thus, we have

$$\hat{\theta_1} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } \hat{\theta_2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Recall that the likelihood function is

$$L(X_i;\theta) = \prod_{i=1}^{n} f(X_i;\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_i-\mu)^2/2\sigma^2} = \frac{1}{(2\pi\sigma^2)^{n/2}} \prod_{i=1}^{n} e^{-(x_i-\mu)^2/2\sigma^2}$$

and so we have our test statistic is

$$\Lambda = \frac{L(\bar{X},\theta_2')}{L(\bar{X},\hat{\theta}_2)} = \frac{\frac{1}{(2\pi\theta_2')^{n/2}} \prod_{i=1}^{n} e^{-(x_i-\bar{X})^2/2\theta_2'}}{\frac{1}{(2\pi\hat{\theta}_2)^{n/2}} \prod_{i=1}^{n} e^{-(x_i-\bar{X})^2/2\hat{\theta}_2}} = (\frac{\hat{\theta}_2}{\theta_2'})^{n/2} \exp(\sum_{i=1}^{n}(x_i-\bar{X})^2(\frac{1}{2\hat{\theta}_2} - \frac{1}{2\sigma_2'}))$$

$$\Lambda = (\frac{\hat{\theta}_2}{\theta_2'})^{n/2} e^{\frac{wn}{2}(\hat{\theta}_2 - \hat{\theta}_2')}$$

where $w = \sum_{i=1}^{n}(X_i-\bar{X})^2/\theta_2'$.

$$\Lambda > k \text{ s.t. } w \le k_1 \text{ or } w \ge k_2 \implies c_1 = \theta_2' k_1 \text{ and } c_2 = \theta_2' k_2$$

6.5.5. Let $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_m$ be independent random samples from the distributions $N(\theta_1,\theta_3)$ and $N(\theta_2,\theta_4)$, respectively.

(a) Show that the likelihood ratio for testing $H_0 : \theta_1 = \theta_2,\ \theta_3 = \theta_4$ against all alternatives is given by

$$\frac{\left[\sum_1^n (x_i-\overline{x})^2/n\right]^{n/2} \left[\sum_1^m (y_i-\overline{y})^2/m\right]^{m/2}}{\left\{\left[\sum_1^n (x_i-u)^2 + \sum_1^m (y_i-u)^2\right] \Big/ (m+n)\right\}^{(n+m)/2}},$$

where $u = (n\overline{x} + m\overline{y})/(n+m)$.

(b) Show that the likelihood ratio test for testing $H_0 : \theta_3 = \theta_4$, θ_1 and θ_2 unspecified, against $H_1 : \theta_3 \neq \theta_4$, θ_1 and θ_2 unspecified, can be based on the random variable

$$F = \frac{\sum_1^n (X_i-\overline{X})^2/(n-1)}{\sum_1^m (Y_i-\overline{Y})^2/(m-1)}.$$

a. On the whole data space Ω, from the given information, we can derive the parameter estimates:

$$\begin{aligned}
\hat{\theta}_1 &= \bar{X}, \\
\hat{\theta}_2 &= \bar{Y}, \\
\hat{\theta}_3 &= \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2, \\
\hat{\theta}_4 &= \frac{1}{m}\sum_{i=1}^{m}(Y_i-\bar{Y})^2.
\end{aligned}$$

Under the null hypothesis H_0

$$\begin{aligned}
\hat{\theta}_{10} &= \hat{\theta}_{20} = u, \\
\hat{\theta}_{30} &= \hat{\theta}_{40} = \frac{1}{n+m}\left[\sum_{i=1}^{n}(X_i - u)^2 + \sum_{i=1}^{m}(Y_i - u)^2\right].
\end{aligned}$$

Using the above parameter estimates, the likelihood ratio Λ can be formed as

$$\Lambda = \frac{L(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4)}{L(\hat{\theta}_{10}, \hat{\theta}_{30})}$$

where L denotes the likelihood function, which, based on the above parameter definitions, gives the previously stated result for Λ.

b. For the entire parameter set, denoted as Ω, the best estimates for these parameters, derived from our data, are

$$\begin{aligned}
\hat{\theta}_1 &= \bar{X}, \\
\hat{\theta}_2 &= \bar{Y}, \\
\hat{\theta}_3 &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2, \\
\hat{\theta}_4 &= \frac{1}{m}\sum_{i=1}^{m}(Y_i - \bar{Y})^2.
\end{aligned}$$

But, when considering the constraints of H_0:

$$\begin{aligned}
\hat{\theta}_{10} &= \bar{X}, \\
\hat{\theta}_{20} &= \bar{Y}, \\
\hat{\theta}_{30} = \hat{\theta}_{40} &= \frac{1}{n+m}\left[\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i=1}^{m}(Y_i - \bar{Y})^2\right].
\end{aligned}$$

Using the likelihood ratio test we assess the probability of observing our data under two scenarios: the restrictions posed by the null hypothesis and without any restrictions. This is computed as:

$$\Lambda = \frac{\text{Probability under } H_0}{\text{Probability without restrictions}}$$

This can be framed in terms of the sample variances:

$$\Lambda = \frac{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{n/2}\left(\sum_{i=1}^{m}(y_i - \bar{y})^2\right)^{m/2}}{\left[\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right) + \left(\sum_{i=1}^{m}(y_i - \bar{y})^2\right)\right]^{(n+m)/2}}$$

Lastly, our F statistic, which represents a ratio of the sample variances, can be related to the LRT. Expressing it as:

$$F = \frac{s_x^2}{s_y^2}$$

From this, it's clear that the LRT statistic can be reformulated in terms of the F statistic, which follows the distribution $F \sim F_{n-1,m-1}$.