

Name: _____

Individual observations can be of types A , B , or C . According to a genetics model, the probabilities of type B and C are equal. A general model states the probabilities of the three types are θ_1 , θ_2 , and θ_3 respectively, where $\theta_1 + \theta_2 + \theta_3 = 1$. An independent sample yields the following data:

A	B	C
63	23	14

Use a likelihood ratio test to test the goodness of fit of the genetics model versus the general model. The following table displays quantiles from several chi-square distributions.

df	0.8	0.9	0.95	0.96	0.97	0.98	0.99	0.995	0.999	0.9995	0.9999
1	1.64	2.71	3.84	4.22	4.71	5.41	6.63	7.88	10.83	12.12	15.14
2	3.22	4.61	5.99	6.44	7.01	7.82	9.21	10.60	13.82	15.20	18.42
3	4.64	6.25	7.81	8.31	8.95	9.84	11.34	12.84	16.27	17.73	21.11

1. (2 points) State formal null and alternative hypotheses using the parameters of the models.

Solution: $H_0: \theta_2 = \theta_3$, $H_1: \theta_2 \neq \theta_3$.

2. (5 points) Find numerical values for the maximum likelihood estimates of the parameters under the null genetics model. *You do not need to show work to derive an expression for the formula of the MLE.*

Solution: Consider combining categories B and C. Then, we know the MLE for the counts in two categories are estimated by observed frequencies, so $\hat{\theta}_1 = 63/100$ and $\hat{\theta}_2 + \hat{\theta}_3 = 37/100$. Since $\theta_2 = \theta_3$ under H_0 , it follows that $\hat{\theta}_2 = \hat{\theta}_3 = 37/200 = 0.185$. This can be found as well by maximizing $63 \log(1 - 2\theta) + 37 \log(\theta)$.

3. (5 points) Find numerical values for the maximum likelihood estimates of the parameters under the general model. *You do not need to show work to derive an expression for the formula of the MLE.*

Solution: The MLEs under the general model are $\hat{\theta}_1 = 0.63$, $\hat{\theta}_2 = 0.23$, and $\hat{\theta}_3 = 0.14$.

4. (5 points) Find a numerical value for the likelihood ratio test statistic.

Solution:

$$-2 \log \Lambda = 2(63 \log(0.63) + 23 \log(0.23) + 14 \log(0.14) - 63 \log(0.63) - 23 \log(0.185) - 14 \log(0.185)) \doteq 2.211$$

5. (2 points) Circle choices or fill in the blanks.

Solution: The p-value for this hypothesis test is equal to the area to the right of 2.211 below the density of a chi-square distribution with 1 degrees of freedom.

6. (1 points) Find a numerical range for the p-value (for example, $0.01 < p < 0.02$) using the table above.

Solution: From the table, the quantile is between 0.8 and 0.9 so $0.1 < p < 0.2$.