Human Genetics:
Problem Set IV

Youngduck Choi
New York University
yc1104@nyu.edu

Abstract

This work contains the solutions to the problem set IV of Human Genetics 2015 course at New York University.

Question 1. Hypothesis Testing I.

Solution. **(a)** Under the null hypothesis of $3 : 1$ segregation, the expected counts of inflated and constricted pods are $1180 \cdot \frac{3}{4} = 885$ and $1180 \cdot \frac{3}{4} = 295$ respectively.

(b) The chi-value for an observed data under a null hypothesis is defined by

$$\chi^2 = \sum \frac{(O - E)^2}{E}.$$

Substituting the given data yields

$$\begin{aligned}\chi^2 &= \frac{(881 - 885)^2}{885} + \frac{(299 - 295)^2}{295} \\ &\approx 0.0723.\end{aligned}$$

Hence, the chi-square value for the observed data under the null hypothesis is approximately 0.0723.

(c) The degrees of freedom in this case is 1. Using the given computational tool, we obtain that the $p-$value is approximately 0.788.

(d) Since the $p-$value is greater than 0.05, we fail to reject the null hypothesis, which stated $3 : 1$ segregation. $\square$

Question 2. Hypothesis Testing II.

Solution. **(a)** $50 : 50$ is a good null hypothesis about the sex ratio of the newborns (girls: boys), because there is a cross between XX and XY, which should be expected to go through uniform and independent segregation from each gender. By drawing the punnett square, with the uniform and independent segregation of XX and XY sex choromosomes, we see that we should expect $50 : 50$ girls and boys.

(b) Under the null hypothesis, the exact probability that all 8 children are girls is $(\frac{1}{2})^8 = \frac{1}{256}$.

(c) We can compute the chi-value in this case to be $\frac{(0-4)^2}{4} + \frac{(8-4)^2}{4}$, which is 8. As the degrees of freedom is 1, the corresponding $p-$value is 0.00468. With a $p-$value threshold of 0.05, we reject the null hypothesis. We, however, expect to reject the null hypothesis, even when it's true, 0.00468% of the time. Therefore, we cannot claim that the null hypothesis is wrong completely. $\square$

Question 3. Probability.

Solution. (a) Since the dice is fair, the probability that we roll a 5 is $\frac{1}{6}$.

(b) First of all, if we roll two dices, there are in total 6^2 different outcomes with respect to the numbers we see. The cases that we see 11 or greater for the sum of two rolls are exactly $(5, 6)$, $(6, 5)$ and $(6, 6)$, where the tuples denote the outcome of the two dices separately. Hence, the exact probability that the total number rolled will be 11 or greater is $\frac{3}{36} = \frac{1}{12}$.

(c) The event that we see the total number 10 or lower is exactly the complement event of the one described in the part (b). Hence, the exact probability that the total number rolled will be 10 or lower is $1 - \frac{1}{12} = \frac{11}{12}$. $\square$