qbs120_ps6_gibran

Gibran Erlangga

10/22/2021

Question 1

(Based on Rice 9.1) A coin is thrown independently 10 times to test the hypothesis that the probability of a heads is $\frac{1}{2}$ vs the alternative that the probability is not $\frac{1}{2}$. The test rejects if the number of observed heads is 0 or 10.

(a) What is the significance level of the test?

From the question above, we know that:

$$H_0 : p = \frac{1}{2}$$
$$H_1 : p \neq \frac{1}{2}$$

With decision function as:
$$d(X) = 1, X \in \{0, 10\}$$
$$d(X) = 0, X \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

Significance level α is defined as the expected value of decision function $d(X)$ under the probability law specified by H_0. Therefore,

$$
\begin{aligned}
\alpha &= E_0(d(X)) \\
&= 1 \cdot P_0(X \in \{0, 10\}) + 0 \cdot P_0(X \in \{1, 2, .., 9\}) \\
&= P_0(X \in \{0, 10\}) \\
&= P_0(X = 0) + P_0(X = 10) \\
&= \binom{10}{0} \cdot p^0 \cdot (1-p)^{10-0} + \binom{10}{10} \cdot p^{10} \cdot (1-p)^{10-10} \\
&= \left(1 - \frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10} \\
&= \frac{2}{2^{10}} \approx 0.00195
\end{aligned}
$$

(b) If the probability of heads is 0.1, what is the power of the test?
$$
\begin{aligned}
1 - \beta &= E_1(d(X)) \\
&= 1 \cdot P_1(X \in \{0, 10\}) + 0 \cdot P_1(X \in \{1, 2, .., 9\}) \\
&= \binom{10}{0} \cdot p^0 \cdot (1-p)^{10-0} + \binom{10}{10} \cdot p^{10} \cdot (1-p)^{10-10} \\
&= (1 - 0.1)^{10} + (0.1)^{10} \approx 0.3487
\end{aligned}
$$

(c) If the test instead rejects if the number of observed heads is 1 or 9, what is the significance level?

$$\alpha = P(X = 0) + P(X = 1) + P(X = 9) + P(X = 10)$$

$$= \binom{10}{0} \cdot p^0 \cdot (1-p)^{10-0} + \binom{10}{1} \cdot p^1 \cdot (1-p)^{10-1} + \binom{10}{9} \cdot p^9 \cdot (1-p)^{10-9} + \binom{10}{10} \cdot p^{10} \cdot (1-p)^{10-10}$$

$$= \binom{10}{0} \cdot 0.5^0 \cdot (0.5)^{10} + \binom{10}{1} \cdot 0.5^1 \cdot (0.5)^9 + \binom{10}{9} \cdot 0.5^9 \cdot (0.5)^1 + \binom{10}{10} \cdot 0.5^{10} \cdot (0.5)^0$$

$$= 0.5^{10} + 10 \cdot 0.5^{10} + 10 \cdot 0.5^{10} + 0.5^{10}$$

$$= 22 \cdot 0.5^{10}$$

$$= 0.02148438$$

Question 2

Based on Rice 9.2) Which of the following hypotheses are simple, and which are composite? Justify your answers.

(a) X follows a uniform distribution on [0,1].
simple, because it gives us the information about the kind of random variable and its parameters.

(b) A die is unbiased.
This is quite tricky - this will be **simple** if its a fair 6-sided die with probability of each side comes as a result of a toss is $\frac{1}{6}$. On the other hand, this will be a **composite** one if it has more than 6 sides (it's possible) and therefore we will have no idea about the random variables involved in the case.

(c) X follows a normal distribution with mean 0 and variance $\sigma^2 > 10$.
composite, since we don't know exactly the value for its variance. It can be $X \sim N(0, 20)$ and $X \sim N(0, 200)$, which have a very different shape.

(d) X follows a normal distribution with mean $\mu = 0$.
composite, as stated in c), we do not know exactly the distribution of the random variable stated in the question, as the variance can be any positive real number.

Question 3

(Based on Rice 9.5) True or false and state why:

(a) The significance level of a statistical test is equal to the probability that the null hypothesis is true.
false - The definition of significance level is the probability of rejecting the null hypothesis, when it is actually true.

(b) If the significance level of a test is decreased, the power would be expected to increase.
false - The power of the test is the probability of rejecting the null hypothesis when it is not true. We are less likely to reject the null hypothesis with low significance level.

(c) If a test is rejected at the significance level α, the probability that the null hypothesis is true equals α.
false - Null hypothesis is not a random variable, so it does not makes sense to talk about probability in this case.

(d) The probability that the null hypothesis is falsely rejected is equal to the power of the test.
false - The power of the test is the ability to correctly reject the null hypothesis.

(e) A type I error occurs when the test statistic falls in the rejection region of the test.
false - type I error is a case where we reject the null hypothesis when it is actually true. However, when the test statistic falls into the rejection region, we reject the null hypothesis, whether it is true or not.

(f) A type II error is more serious than a type I error.
false - Type I error happens when we mistakenly reject the null hypothesis when it is actually true, while the type II error occurs when we do not reject the null hypothesis when the alternate hypothesis is true. Type II error is not necessarily more serious than type I error, as it highly depends on case at hand. For instance, imagine you got tested for COVID and we have two hypotheses as the outcome:
H_0 : you have COVID, and
H_A : you do not have COVID
In this case, type II error is more important to be addressed as the consequences of it is that you will be diagnosed that you do not have COVID, where in fact you do have one.

(g) The power of a test is determined by the null distribution of the test statistic.
false - The power of the test is the probability of rejecting the null hypothesis when in fact it is not true. However, the null hypothesis has no effect in determining the power of the test.

(h) The likelihood ratio is a random variable.
true - the likelihood ratio is the ratio of likelihood of the observed sample under the null hypothesis to the likelihood of the observed sample under the alternate hypothesis. Different samples will produce different likelihood ratios, so it is a random variable.

Question 4

(Based on Rice 9.30) Suppose that the null hypothesis is true, that the distribution of the test statistic, T say, is continuous with cdf F and that the test rejects for large values of T. Let V denote the p-value of the test.

(a) Show that V = 1 −F(T).
We know that:

1) null hypothesis is true,
2) T is a continuous RV with CDF F,
3) the test rejects for large values of T,
4) V = p-value of the test

say $T = t$ for the observed value of the test statistic, then the p-value of the test become:

$$V = P(T > t|H_0)$$
$$= 1 - P(T \leq t|H_0)$$
$$= 1 - F(t)$$

(b) Conclude that the null distribution of V is uniform. For $x \in [0, 1]$, the cdf of V defined as:

$$F_V(x) = P(V \leq x)$$
$$= P(1 - F(T) \leq x)$$
$$= P(F(T) \geq 1 - x)$$
$$= 1 - P(F(T) < 1 - x)$$
$$= 1 - (1 - x)$$
$$= x$$

3

Applying derivative to the result, we observe that the density function of V is:

$$F_V'(x) = (x)' = 1, x \in [0, 1]$$

Which is exactly the density function of a uniform random variable.

(c) If the null hypothesis is true, what is the probability that the p-value is greater than .1?
from b) we know that the p-value is a uniform random variable on $[0, 1]$. Therefore,

$$P(V > 0.1) = P(1 - V \leq 0.1)$$
$$= 1 - F_V(0.1)$$
$$= 1 - 0.1$$
$$= 0.9$$

(d) Show that the test that rejects if $V < \alpha$ has significance level α.
We know that V is uniform and the significance level of a test is the expected value of the decision function under the probability law described in the null hypothesis, so:

$$E(d(X|H_0)) = P(V < \alpha|H_0)$$
$$= F_V(\alpha) = \alpha$$

Question 5

(Based on Rice 9.42) Nylon bars were tested for brittleness. Each of the 280 bars was molded under similar conditions and was tested in five places. Assuming that each bar has uniform composition, the number of breaks on a given bar should be binomially distributed with five trials and an unknown probability of p of failure. If the bars are all of the same uniform strength, p should be the same for all of them; if they are of different strengths, p should vary from bar to bar. Thus, the H0 is that the p's are all equal. The following table summarizes the outcome of the experiment:

(a) Under the given assumption, the data in the table consist of 280 observations of independent binomial random variables. Find the mle of p.

Let p_i be the probability that the i-th bar breaks for every $i \in \{1, 2, ..., 280\}$. Denoting the X_i as the total number of bar breaks with the i-th bar, then X_i is a binomial RV with $n = 5$ represents the number of trials and p_i represents the probability of success. Hence, the probability that the i-th bar breaks at k places becomes:

$$P(X_i = k) = \binom{5}{k} \cdot p_i^k \cdot (1 - p_i)^{5-k}, k \in \{0, 1, 2, 3, 4, 5\}$$

the null hypothesis that all bars p_i are equal to p, hence:

$$\theta_k(p) = \binom{5}{k} \cdot p^k \cdot (1 - p)^{5-k}, k \in \{0, 1, 2, 3, 4, 5\}$$

which defined as the probability that any bar breaks at k places.

Let O_k be the observed counts, i.e. O_0 is the number of bars that did not break in any of the five places, and so on).

4

Under the null hypothesis, the likelihood function is:

$$lik(p) = \prod_{k=0}^{5} \theta_k(p)^{O_i}$$

$$= \left[\prod_{k=0}^{5} \binom{5}{k} \right] \cdot p^{\sum_{k=0}^{5} k \cdot O_k} \cdot (1-p)^{\sum_{k=0}^{5} (5-k) \cdot O_k}$$

Applying natural logarithm on both ends for easier calculation, we get:

$$l(p) = ln(lik(p)) = \sum_{k=0}^{5} ln\left[\binom{5}{k} \right] + \sum_{k=0}^{5} k \cdot O_k \cdot ln(p) + \sum_{k=0}^{5} (5-k) \cdot O_k \cdot ln(1-p)$$

Derivative of l with respect to p is:

$$l'(p) = \frac{1}{p} \cdot \sum_{k=0}^{5} k \cdot O_k - \frac{1}{1-p} \cdot \sum_{k=0}^{5} (5-k) \cdot O_k$$

Stationary points are the null-points of the derivative, hence:

$$l'(p) = 0 \Leftrightarrow \frac{1}{p} \cdot \sum_{k=0}^{5} k \cdot O_k - \frac{1}{1-p} \cdot \sum_{k=0}^{5} (5-k) \cdot O_k = 0$$

$$\Leftrightarrow (1-p) \cdot \sum_{k=0}^{5} k \cdot O_k - p \cdot \sum_{k=0}^{5} (5-k) \cdot O_k = 0$$

$$\Leftrightarrow p = \frac{\sum_{k=0}^{5} k \cdot O_k}{5 \sum_{k=0}^{5} O_k}$$

Calculating the second derivative of l with respect to p, we get:

$$l''(p) = \frac{1}{p^2} \sum_{k=0}^{5} k \cdot O_k - \frac{1}{(1-p)^2} \sum_{k=0}^{5} (5-k) \cdot O_k < 0$$

Therefore, we know that the log-likelihood function is concave. Hence:

$$\hat{p} = \frac{\sum_{k=0}^{5} k \cdot O_k}{5 \sum_{k=0}^{5} O_k}$$

$$= \frac{199}{5 \cdot 280}$$

$$= 0.142$$

(b) Pooling the last three cells, test the agreement of the observed frequency distribution with the binomial distribution using Pearson's chi-square test.

5

```r
get_theta <- function(n, k, p) {
  return(round(choose(n, k) * p^k * (1-p)^(n-k), digits=7))
}

theta_0 <- get_theta(5, 0, 0.142)
theta_1 <- get_theta(5, 1, 0.142)
theta_2 <- get_theta(5, 2, 0.142)
theta_3 <- get_theta(5, 3, 0.142)
theta_4 <- get_theta(5, 4, 0.142)
theta_5 <- get_theta(5, 5, 0.142)

cat(paste("theta value for k = 0 is", theta_0, "\n",
          "theta value for k = 1 is", theta_1, "\n",
          "theta value for k = 2 is", theta_2, "\n",
          "theta value for k = 3 is", theta_3, "\n",
          "theta value for k = 4 is", theta_4, "\n",
          "theta value for k = 5 is", theta_5, "\n"
          ))
```

```
## theta value for k = 0 is 0.4649823
##  theta value for k = 1 is 0.3847756
##  theta value for k = 2 is 0.1273616
##  theta value for k = 3 is 0.0210785
##  theta value for k = 4 is 0.0017443
##  theta value for k = 5 is 5.77e-05
```

We want to test:
$$H_0 : \theta_k(p) = \theta_k(\widehat{p})$$

Using Pearson's χ^2 statistic,

$$\chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$$

with $m = 6$, $k = 1$ -> distributed to a χ^2 distribution with 4 degree of freedom.

```r
get_pearson <- function(a, b) {
  return((a-b)^2 / b)
}

bar <- c(0, 1, 2, 3, 4, 5)
freq <- c(157, 69, 35, 17, 1, 1)
e <- 280 * c(theta_0, theta_1, theta_2, theta_3, theta_4, theta_5)

pearson_val <- c()

chi1 <- (freq[1]-e[1])^2 / e[1]
chi2 <- (freq[2]-e[2])^2 / e[2]
chi3 <- (freq[3]-e[3])^2 / e[3]
chi4 <- ((freq[4]+freq[5]+freq[6])-(e[4]+e[5]+e[6]))^2 / (e[4]+e[5]+e[6])

chi <- chi1 + chi2 + chi3 + chi4
(res <- 1 - pchisq(chi,2))
```

```
## [1] 2.506155e-10
```

Question 6

In this problem you will develop and apply a test for normality based on the coefficient of skewness as defined in Section 9.9 of Rice. (a) Write an R function to compute the coefficient of skewness, b1.

```
get_coef_skewness <- function(x) {
  return(sum((x - mean(x)^3)/var(x)^3)/length(x))
}
```

(b) Use simulation to approximate the sampling distribution of b1 when the data is modeled by 100 independent N(0,1) RVs.

```
b1 <- c()

for (i in 1:1000) {
  val <- get_coef_skewness(rnorm(100))
  b1 <- append(b1, val)
}

plot(density(b1))
```

density.default(x = b1)

N = 1000   Bandwidth = 0.02469
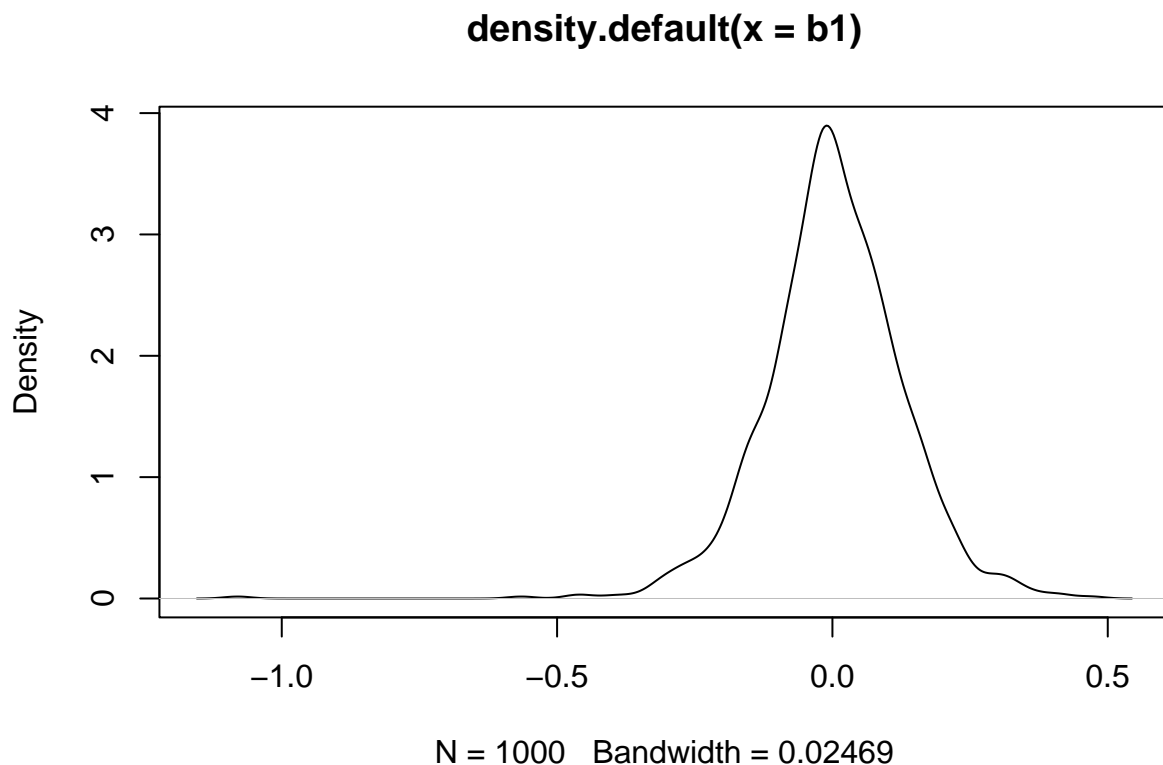
(c) Test 1000 data sets each containing 100 independent N(0,1) RVs for normality. For this test H0 is that the data are iid N(0,1) and HAis that the data are not iid N(0,1). Use the sampling distribution of b1 to compute an approximate p-value for each test. Generate a Q-Q plot of these p-values relative to a theoretical U(0,1) distribution. Does the shape of this plot match your expectations? Explain.

7

```r
test <- c()
p_value <- c()

for (i in 1:1000) {
  val <- get_coef_skewness(rnorm(100))
  test <- append(test, val)
}

for (i in 1:1000) {
  lower_bound <- sum(b1 < test[i])/length(b1)
  upper_bound <- sum(b1 > test[i])/length(b1)
  res <- min(lower_bound, upper_bound)*2
  p_value <- append(p_value, res)
}

qqplot(qunif(seq(0, 1, 0.001)), p_value)
```
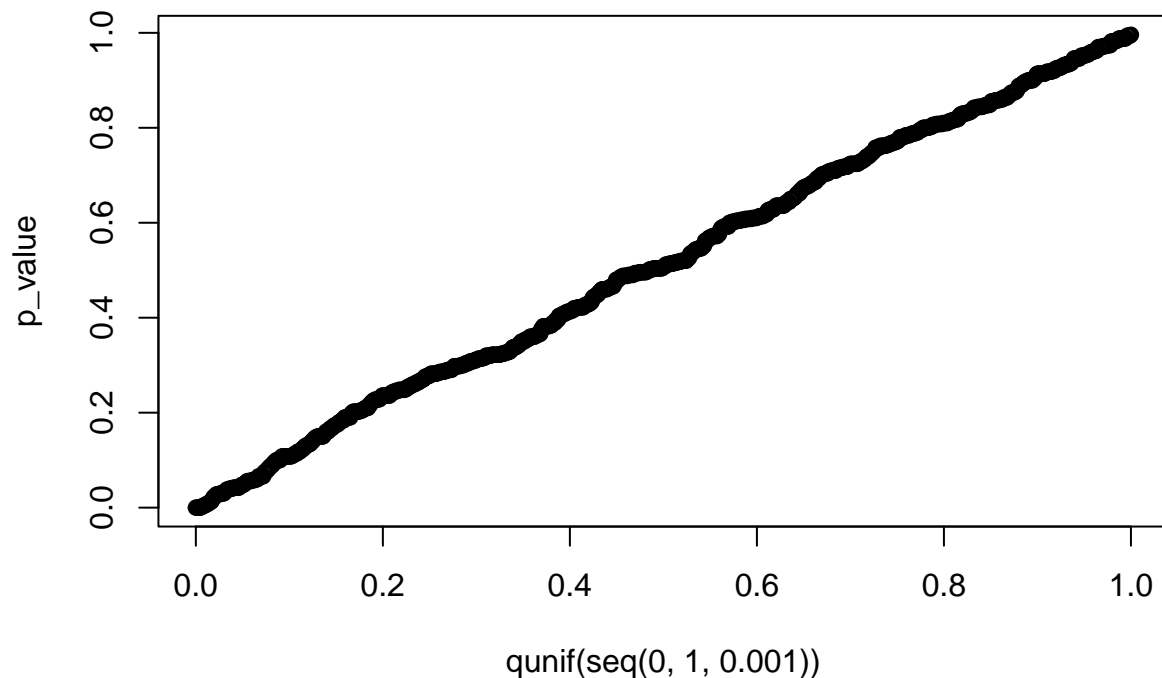

(d) Test 1000 data sets each containing 100 independent Poisson RVs with $\lambda = 1$ for nor- mality. Compute an approximate p-value for each test. Generate a Q-Q plot of these p-values relative to a theoretical U(0,1) distribution. Does the shape of this plot match your expectations? Explain.

```r
test <- c()
p_value <- c()

for (i in 1:1000) {
```

```
  val <- get_coef_skewness(rpois(100, 1))
  test <- c(test, val)
}

for (i in 1:1000) {
  lower_bound <- sum(b1 < test[i])/length(b1)
  upper_bound <- sum(b1 > test[i])/length(b1)
  res <- min(lower_bound, upper_bound)*2
  p_value <- c(p_value, res)
}

qqplot(qunif(seq(0, 1, 0.001)), p_value)
```