STATISTICS
Homework 9

Due June 25, 2021

1. A random sample from a continuous uniform distribution $U[0,\ \theta]$ is given in the following table. A hypothesis test $H_0: \theta = 5,\quad H_0: \theta \neq 5$ is to be conducted.

0.9008	3.5321	4.9546	2.9965	4.6778
2.2499	3.4768	0.295	3.8101	3.7739
2.7089	2.7323	0.1799	2.1793	2.5436
1.545	4.8106	1.4242	2.0554	4.5775
4.6189	2.3098	1.8262	1.2262	4.8855

   (1) Conduct the test using $T = \sum_{i=1}^{n}\left[-\ln\left(\frac{X_i}{5}\right)\right]$ as a test statistic and level of significance $\alpha = 0.05$.

   (2) Calculate and plot the power function of the test for $\theta = 2.0,\ 2.1, \cdots,\ 9.9,\ 10$.

2. Use the same data and hypotheses in Problem 1.

   (1) Conduct the test using $T = \frac{Y_n}{5}$ and level of significance $\alpha = 0.05$.

   (2) Calculate and plot the power function of the test for $\theta = 2.0,\ 2.1, \cdots,\ 9.9,\ 10$.

3. Final grades of several students in class are listed in the following table.

69	46	51	34	85	79	49	58	52	57	50
39	74	46	57	45	49	44	60	53	59	24
75	30	70	44	20	34	84	45	55	57	53
54	35	38	49	52	45	79	68	63	57	65

   At level of significance $\alpha = 0.05$, test whether the grade distribution is significantly different from a normal distribution with mean $\mu = 60$ and standard deviation $\sigma = 15$, using the ***chisq.test*** function in R. Set the number of categories equal to 6.

4. The Kolmogorov-Smirnov GOF test can be implemented using ***ks.test*** in R. The ks.test is rather straightforward for one-sample case. A brief description is as follows:

   ks.test(x, y, parameters, alternative=”two.sided”)

   where *x* is the data vector to be tested, *y* is a string vector specifying the hypothesized distribution, *parameters* are the values of distribution parameters corresponding to *y*, and *alternative* represents a string vector (“*less*”, “*greater*”, or “*two.sided*”) for one-tail or two-tail test.

   Examples:

   ks.test(x, ”pnorm”, 30, 10, alternative=”two.sided”)

   ks.test(x, ”pexp”, 0.2, alternative=”greater”)

   At level of significance $\alpha = 0.05$, test whether data in the following table is significantly different from a normal distribution with mean 60 and standard deviation 15.

51.29	53.45	75.82	49.08	81.47
49.38	58.61	66.07	65.27	55.49
43.85	44.04	57.11	63.92	80.54
42.82	72.91	78.42	71.39	81.40

5. The following table shows the number of home runs in a game. Assuming that the distribution of home runs is the same for each game and independent between games, conduct the following hypotheses test:

H_0 : The home runs can be modelled by a Poisson distribution.
H_1 : The home runs cannot be modelled by a Poisson distribution.

Number of home runs	**Number of games**
0	**43**
1	**52**
2	**40**
3	**17**
4	**9**
5	**1**