Composite Hypotheses

Homework 7

Problems

1. Let $X_1, \ldots, X_n \sim Ber(p)$ be n Bernoulli trials and consider the test

$$H_0 : p \leq p_0 \quad \text{versus} \quad H_1 : p > p_0.$$

   (a) Then $T(\mathbf{x}) = \sum_{i=1}^{n} x_i$ is a minimal sufficient statistic. Show that the choice of a rejection region

$$\{\mathbf{x}; T(\mathbf{x}) \geq \tilde{k}_\alpha\}$$

   is a uniformly most powerful test.

   (b) What is the distribution of $T(\mathbf{x})$.

   (c) For $n = 20$, $p_0 = 0.6$ and $\alpha = 0.05$ and $\alpha = 0.01$, what is $\tilde{k}_\alpha$?

   (d) for what values of α is

$$\{\mathbf{x}; T(\mathbf{x}) \geq 15\}$$

   a critical region?

2. Let $X_1, \ldots, X_m \sim Unif(0, \theta)$ and consider the test

$$H_0 : \theta \leq 1 \quad \text{versus} \quad H_1 : \theta > 1.$$

   (a) Let $T(\mathbf{X}) = \max_{1 \leq i \leq n} X_i$ and find k_α so that the rejection region

$$\{\mathbf{x}; T(\mathbf{x}) \geq k_\alpha\}$$

   has probability α for the choice $\theta = 1$.

   (b) For $m = 5$ and $T(\mathbf{X}) = 0.99$, find the p-value.

   (c) For $m = 5$ and $T(\mathbf{X}) = 1.01$, find the p-value.

   (d) Determine the power function and provide a plot in the case $m = 5$ and $\alpha = 0.05$. Indicate $k_{0.05}$ on the plot.

3. For a two-samples t test,

$$H_0 : \mu_1 \leq \mu_2 \quad \text{versus} \quad H_1 : \mu_2 > \mu_1,$$

   a matched pair procedure is used when the two pairs observations $X_{1,1}, \ldots, X_{1,m}$ and $X_{2,1}, \ldots, X_{2,m}$ are made on the same individuals. In this case the individual observations are subtracted and the t-test is performed on the differences,

$$X_i = X_{2,i} - X_{1,i}.$$

(a) July and August are the strongest months for monsoon rains in Tucson. Your friend suspects that more rain takes place in July. Give an appropriate hypothesis for this situation,

(b) For the years 2007 to 2016, here are the data for rainfall total in inches.

	2007	2008	2009	2010	2011	2012	2013	2104	2015	2016
July	4.38	4.03	2.15	1.44	1.69	0.71	1.44	4.58	2.06	3.32
August	4.27	1.69	0.86	2.89	1.03	0.00	1.85	3.32	1.80	1.09

Perform a match pair t-test on these data and report the p-value.

Challenging Problems

1. For a two-samples t test,
$$H_0 : \mu_1 \leq \mu_2 \quad \text{versus} \quad H_1 : \mu_2 > \mu_1,$$
the two-sample procedure uses the test statistic
$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}.$$
   (a) What is the test statistic under the two-sample procedure for monsoon data above?
   (b) Which of the two test statistics is larger? Explain why.

2. Consider the hypothesis for exponential random variables with rate parameter β.
$$H_0 : \beta \geq 1 \quad \text{versus} \quad H_1 : \beta < 1,$$
   (a) What is the distribution of the test statistic $T(\mathbf{x}) = \sum_{i=1}^{n} x_i$.
   (b) For $n = 20$, give the critical value for an $\alpha = 0.05$ level test.
   (c) What is the power of the test?
   (d) Draw the receiving operator characteristic for $\beta = 1$ against the alternative $\beta = 1/2$.
   (e) Simulate data for 20 observations with $\beta = 1$ followed by 5 observations for $\beta = 1/2$. Use the Fisher test and compute the p-values for the first m tests where m ranges from 1 to 25.