SM-4331 Exercise 4

1. Suppose that $X_1, ..., X_n$ and $Y_1, ..., Y_n$ are two independent random samples from two exponential distributions with mean μ_1 and μ_2 respectively.
   (a) Find the likelihood ratio test for statistic $H_0 : \mu_1 = \mu_2$ against $H_0 : \mu_1 \neq \mu_2$.
   (b) Specify the asymptotic distribution of the test statistic under H_0.

2. A survey of the use a particular product was conducted in four areas, and a random sample of 200 potential users was interviewed in each area. In area i, for $i = 1, 2, 3, 4$, X_i of the 200 said that they used the product. Construct a likelihood ratio test to test whether the proportion of the population using the product is the same in each of the four areas. Carry out the test at 5% level for the case $X_1 = 76$, $X_2 = 53$, $X_3 = 59$ and $X_4 = 48$.

3. Let $\mathbf{X} = \{X_1, \ldots, X_n\}$ be an independent sample from a distribution with pdf

$$\mathrm{P}(X_i = x) = f(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

   (a) Write down the log-likelihood function and the score function, and then obtain the ML estimate for λ.
   (b) Determine the Fisher information for λ.
   (c) Suppose that we observe $\bar{X} = 3.5$ from $n = 10$ observations. Test at the 5% significance level that

$$H_0 : \lambda = 5 \quad \text{v.s.} \quad H_1 : \lambda \neq 5$$

      i. Using the Wald test.
      ii. Using the Score test.

4. In 1861, 10 essays appeared in the New Orleans Daily Crescent. They were signed "Quintus Curtius Snodgrass" and some people suspected they were actually written by Mark Twain. To investigate this, we will consider the proportion of three letter words found in an authors work. From 8 of Twains essays, the proportions are:

   0.225 0.262 0.217 0.240 0.230 0.229 0.235 0.217

   From 10 Snodgrass essays, the proportions are:

   0.209 0.205 0.196 0.210 0.202 0.207 0.224 0.223 0.220 0.201

   Perform a Wald test for equality of the means. Report the p-value and a 95% confidence interval for the difference of means. What do you conclude?

5. A sample of 11 observations from population $\mathrm{N}(\mu, \sigma^2)$ yields the sample mean $\bar{X} = 8.68$ and the sample variance $s^2 = 1.21$. At 5% significance level, test the following hypotheses.

(a) $H_0 : \mu = 8$ against $H_1 : \mu > 8$

(b) $H_0 : \mu = 8$ against $H_1 : \mu < 8$

(c) $H_0 : \mu = 8$ against $H_1 : \mu \neq 8$

Repeat the above exercise with the additional assumption $\sigma^2 = 1.21$. Compare the results with those derived without this assumption and comment.

6. There is a theory that people can postpone their death until after an important event. To test this theory, Phillips and King (1988, Lancet, pp. 728-) collected data on deaths around the Jewish holiday Passover. Of 1919 deaths, 922 died the week before the holiday and 997 died the week after. Think of this as a binomial and test the null hypothesis that $\theta = 1/2$, where θ is the probability that a death occurs after the holiday. Also construct a confidence interval for θ.

7. (a) Two independent random samples, of n_1 and n_2 observations, are drawn from normal distributions with the same variance σ^2. Let s_1^2 and s_2^2 be the sample variances of the first and the second sample, respectively. Show that

$$\hat{\sigma}^2 = \frac{1}{n_1 + n_2 - 2}\left\{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right\}$$

is an unbiased estimator for σ.

(b) Two makes of car safety belts, A and B have breaking strengths which are normally distributed with the same variance. A sample of 140 belts of make A and a sample of 220 belts of make B were tested. The sample means and the sums of squares about the means (i.e. $\sum_i (X_i - \bar{X})^2$) of of the breaking strengths (in lbf units) were $(2685, 19000)$ for make A, and $(2680, 34000)$ for make B. Is there any significant evidence to support the hypothesis that belts of make A are stronger than belts of make B?

8. The table below summarized the fate of the passengers and the crew when the Titanic sank on Monday, 15 April 1912. Test the hypothesis of independence between the row variable and the column variable in the table, and interpret your findings.

	Men	Women	Boys	Girl
Survived	332	318	29	27
Died	1360	104	35	18