Stat 513
Emerson, Win 16
Homework #7
March 4, 2016

**Written solutions to the homework problems are due on Friday, March 11, 2016 at the beginning of class.**

**As noted on the syllabus, copying of homework solutions is not allowed and, when detected, will be investigated as an infraction of the academy integrity policy of the University of Washington. While it is permissible to discuss problems with other students, TAs, or the instructor in order to learn how to solve a problem, your written solutions must be prepared without directly referencing any notes or solutions derived from other students or sources found on the internet.**

## REGULAR PROBLEMS

1. Suppose that $Y_1, Y_2, \ldots, Y_n$ are i.i.d. with density
$$f_Y = e^{-(y-\mu)}\mathbf{1}_{(\mu,\infty)}$$
for $\mu \in (-\infty, \infty)$. Use Basu's theorem to show that $Y_{(1)}$ is independent of the sample variance $s^2$.

2. Suppose that for data $Y_1, Y_2, \ldots, Y_n$, conditional on the parameter $p$, have distribution as i.i.d. Bernoulli random variables with $Y_i|p \sim \mathcal{B}(1, p)$, $p \in (0, 1)$. Further suppose prior distribution $p \sim \ Beta(\alpha, \beta)$.
   (a) Show that this prior distribution is the conjugate prior distribution.
   (b) Find the posterior distribution of $p$ conditional on $\vec{Y} = (Y_1, \ldots, Y_n)^T$.
   (c) Find the Bayes estimator of $p$ based on squared error loss. Can this estimator be expressed as a weighted average of the data and the prior distribution? If so, do so.
   (d) Can you find a closed form solution for the Bayes estimator of $p$ based on absolute error loss. If so, do so. If not, explain why not.

3. Suppose that for data $Y_1, Y_2, \ldots, Y_n$, conditional on the hazard parameter $\lambda$, have distribution as i.i.d. exponential random variables with $Y_i|p \sim \mathcal{E}(\lambda)$, $\lambda \in (0, \infty)$. Further suppose prior distribution $\lambda \sim \ Gamma(\alpha, \beta)$.
   (a) Show that this prior distribution is the conjugate prior distribution.
   (b) Find the posterior distribution of $\lambda$ conditional on $\vec{Y} = (Y_1, \ldots, Y_n)^T$.
   (c) Find the Bayes estimator of $\lambda$ based on squared error loss. Can this estimator be expressed as a weighted average of the data and the prior distribution? If so, do so.

(d) Can you find a closed form solution for the Bayes estimator of $\lambda$ based on absolute error loss. If so, do so. If not, explain why not.

4. Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. gamma random variables with $Y_i \sim \Gamma(\gamma, \lambda)$ with $\gamma > 0$ known and $\lambda > 0$ the rate parameter (so $E[Y_i] = \gamma/\lambda$.

   (a) Find the most powerful level $\alpha$ test of the simple hypotheses $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda = \lambda_1 < \lambda_0$. Show that the test is a function of the sufficient statistic, and explicitly provide a definition of the critical value. Is the test unbiased? Is the test consistent?

   (b) Extend the test of part a to the test of the composite hypotheses $H_0 : \lambda \geq \lambda_0$ versus $H_1 : \lambda = \lambda_1 < \lambda_0$. Show that the power function is decreasing in $\lambda$. Is the test unbiased? Is the test consistent?

   (c) Define a two-sided $100(1-\alpha)\%$ confidence interval for $\lambda$.

5. Suppose $Y_1, Y_2, \ldots, Y_n$ are independent random variables with $Y_i \sim (\mu_i, \sigma^2)$ and the distribution is otherwise unspecified. Further suppose that for known $x_1, \ldots, x_n$, $\mu_i = \beta_0 + \beta_1 x_i$. We are interested in testing $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 \neq 0$.

   (a) First presuming that $\sigma^2$ is known, derive an approximate level $\alpha$ test of $H_0$ versus $H_1$, providing an explicit expression for the critical value. Be sure to explain the conditions under which your test will be a good approximation to a level $\alpha$ test.

   (b) Show that your test in part a corresponds to a likelihood ratio test (LRT) when the $Y_i$'s satisfy $Y_i | x_i \sim \mathcal{N}(\mu_i, \sigma^2)$.

   (c) Now modify your test in part a to the case where $\sigma^2$ is unknown.

   (d) How would your test in part c differ from a test derived when the data were known to be normally distributed? Are the tests asymptotically equivalent?

   (e) Find an expression for a

6. Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. Bernoulli random variables with $Y_i \sim \mathcal{B}(1, p)$, $p \in (0, 1)$.

   (a) Derive an expression for the upper $100(1-\alpha)\%$ confidence bound for $p$ using the exact distribution of $\vec{Y}$.

   (b) Suppose we observe $Y_i = 0$ for all $i = 1, \ldots, n$. Show that for large $n$, the bound found in part a is approximately $3/n$ when $\alpha = .05$ and approximately $3.69/n$ when $\alpha = .025$

## MORE INVOLVED PROBLEMS

7. Suppose $Y_1, Y_2, \ldots, Y_{n_Y}$ are i.i.d. Bernoulli random variables with $Y_i \sim \mathcal{B}(1, p_Y)$ and $X_1, X_2, \ldots, X_{n_X}$ are i.i.d. Bernoulli random variables with $X_i \sim \mathcal{B}(1, p_X)$, with $p_Y, p_X \in (0, 1)$. Our target of inference is $\theta = p_y - p_x$, interested in approximate two-sided level $\alpha$ tests of $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$ and $100(1-\alpha)\%$ confidence intervals for $\theta$.

(a) Provide an expression for the Wald test (including critical value).

(b) Provide an expression for the score test (including critical value).

(c) Provide an expression for the likelihood ratio test (including critical value).

(d) Provide an expression for a CI based on the Wald test.

(e) Describe the steps involved in finding CI based on score or LRT.

8. (optional) Suppose we observe data $Y_1, Y_2, \ldots, Y_n$ that are i.i.d. nonnegative random variables with $Y_i \sim (\theta, \gamma\theta^2)$. Our target of interest is to decide whether $Y_i$ is distributed according to a gamma distribution or a lognormal distribution. That is we want a most powerful level $\alpha$ test of $H_0$ vs $H_1$ with

$$H_0 : Y \sim f_Y(y \mid \eta, \gamma) = \frac{1}{\Gamma(\eta)\gamma^\eta} y^{\eta-1} e^{-y/\gamma} \qquad \text{for } \eta = \theta/\gamma$$

$$H_1 : Y \sim f_Y(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma y} \exp\left\{ -\frac{(\log(y) - \mu)^2}{2\sigma^2} \right\} \qquad \text{for } \begin{cases} \sigma^2 = \log(\gamma + 1) \\ \eta = \log(\theta) - \frac{\sigma^2}{2} \end{cases}$$

(a) Suppose it is known $\mu = 0$ and $\eta = 1$ (so $\gamma = (1 + \sqrt{(}5))/2$ and $\sigma^2 = 2\log(\gamma)$). Derive the MP-$\alpha$ test, describing how the critical value could be found.

(b) Provide an expression for the MP-$\alpha$ test when all distributional parameters are unknown.