# STA 114: Statistics

## Practice Problems 2

1. For each question below <u>record</u> which stated option is correct and <u>give</u> a very brief explanation why.

   (a) Jeffreys prior is improper for every statistical model.
   A. True  B. False

   (b) If $X = (X_1, \cdots, X_n)$ and $X_{n+1}$ are modeled as $X_i \overset{\text{IID}}{\sim} \mathsf{Poisson}(\mu)$, then for any prior pdf $\xi(\mu)$, $P(X_{n+1} = 0 \mid X = x) = \int_0^\infty e^{-\mu}\xi(\mu \mid x) d\mu$.
   A. True  B. False

   (c) Let $X \sim \mathsf{Binomial}(n, p)$, with $p \in [0, 1]$ unknown. Among the two families, $\{\mathsf{Beta}(a, b) : a \geq 1, b \geq 1\}$ and $\{\mathsf{Beta}(a, b) : 0 < a < 1, 0 < b < 1\}$, a conjugate prior family for $p$ is found in
   A. the first only  B. the second only  C. both  D. neither

   (d) For $X = (X_1, \cdots, X_n)$ with $X_i \overset{\text{IID}}{\sim} \mathsf{Uniform}(0, \theta)$, $\theta > 0$. Suppose $\theta$ is assigned a prior pdf $\xi(\theta)$ such that $\xi(\theta) > 0$ at every $\theta > 0$. . Then $P(\theta < x_{\max} \mid X = x) = 0$ for any observed data $x = (x_1, \cdots, x_n)$, where $x_{\max} = \max(x_1, \cdots, x_n)$.
   A. True  B. False

2. Consider data $X = (X_1, \cdots, X_n)$ consisting of $n = 15$ observations. Below are two possible statistical models for this data and a prior pdf on parameters for each model. For each setting <u>calculate</u> the posterior probability of the stated event if we observe $X = x$ with $\bar{x} = 3.33$, $\sum_{i=1}^n (x_i - \bar{x})^2 = 55.73$.

   (a) *Model.* $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$ with $-\infty < \mu < \infty$ unknown and $\sigma^2 = 2.5$.
   *Prior.* $\xi(\mu) = \text{const.}$, $-\infty < \mu < \infty$.
   *Event.* $\mu > 2.5$.

   (b) *Model.* $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$ with $-\infty < \mu < \infty$ and $\sigma^2 > 0$ unknown.
   *Prior.* $\xi(\mu, \sigma^2) = \text{const.}/\sigma^2$, $-\infty < \mu < \infty$, $\sigma^2 > 0$.
   *Event.* $\mu > 2.5$.

3. Suppose a positive, integer valued random variable $X$ is modeled as $X \sim f(x \mid p)$, $p \in [0, 1]$ where the pmf $f(x \mid p)$ is given by,

$$f(x \mid p) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad \text{if } x = k, k+1, k+2, \cdots$$

and $f(x \mid p) = 0$ otherwise [This is known as the negative binomial pmf]. Here, $k$ is a fixed, known integer and $p \in [0, 1]$ is an unknown parameter. <u>Identify</u> a conjugate prior family for $p$ (among the well known, named distributions) and <u>EXPLAIN</u> in detail why it is conjugate.

4. For data $X = (X_1, X_2, \cdots, X_n)$, consider the model $X_i \overset{\text{IID}}{\sim} \mathsf{IG}(\mu, \lambda)$, i.e., the pdf of each $X_i$ is given by the inverse-Gaussian pdf: $(\frac{\lambda}{2\pi x_i^3})^{1/2} \exp\{-\frac{\lambda(x_i-\mu)^2}{2\mu^2 x_i}\}$ if $x_i > 0$ and 0 if $x_i \leq 0$. Here $\lambda > 0$ is assumed known and $\mu > 0$ is an unknown parameter. <u>FIND</u> the Jeffreys' prior for $\mu$. [Hint: if $Y \sim \mathsf{IG}(\mu, \lambda)$ then $\text{E}Y = \mu$.]

5. A machine goes through 4 hazard levels $\theta$, coded 0 through 3 (from low hazard to high hazard) with use over time. The hazard level of a day can be measured by the frequency of hazardous incidents $X$ on that day, again coded 0 through 3 (low frequency to high frequency). Suppose $X$ is modeled with pmfs $f(x|\theta)$, $\theta \in \Theta = \{0, 1, 2, 3\}$ as given by the rows of the following table.

| $\theta$ | $f(0\|\theta)$ | $f(1\|\theta)$ | $f(2\|\theta)$ | $f(3\|\theta)$ |
|---|---|---|---|---|
| 0 | $\frac{4}{10}$ | $\frac{3}{10}$ | $\frac{2}{10}$ | $\frac{1}{10}$ |
| 1 | 0 | $\frac{3}{6}$ | $\frac{2}{6}$ | $\frac{1}{6}$ |
| 2 | 0 | 0 | $\frac{2}{3}$ | $\frac{1}{3}$ |
| 3 | 0 | 0 | 0 | 1 |

Frequencies of hazardous incidents from different days are assumed to be conditionally independent of each other given the underlying hazard levels of those days.

(a) Suppose $X$ denotes today's measurement of the machine's frequency of hazardous incident, while $X^*$ denotes the same for tomorrow. Suppose the hazard level today, $\theta$, is assigned a discrete uniform prior pmf $\xi(0) = \xi(1) = \xi(2) = \xi(3) = 1/4$. Assuming the hazard level does not change from today to tomorrow, calculate the posterior predictive probabilities that $X^*$ is going to be at least as much as $X$, for each of the 4 possible observations $X = 0$, $X = 1$, $X = 2$ and $X = 3$.

(b) Now suppose $X$ is as before, but $X^*$ denotes the frequency of hazardous incidents measured on the day exactly a month from now. Assuming the hazard level goes up one unit in a month's time (unless it is already 3, in which case it remains at 3), calculate the posterior predictive probabilities that $X^*$ will be at least as much as $X$ for each of the 4 possible observations $X = 0$, $X = 1$, $X = 2$ and $X = 3$. [Still using the discrete uniform prior for today's hazard level.]

(c) Repeat parts (a) and (b), but use MLE plug-in to calculate the predictive probabilities.

6. Two sets of data $X = (X_1, \cdots, X_n)$ and $Y = (Y_1, \cdots, Y_m)$ are modeled as $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_1, \sigma^2)$, $Y_j \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_2, \sigma^2)$, $X_i$'s and $Y_j$'s independent, with $-\infty < \mu_1, \mu_2 < \infty$ and $\sigma^2 > 0$ unknown. According to an yet untested theory (call it $H_+$) $\mu_1$ must be larger than $\mu_2$. We know that for this model a $100(1-\alpha)\%$ ML confidence interval for

$\eta = \mu_1 - \mu_2$, based on observations $(x, y)$ is given by

$$(\bar{x} - \bar{y}) \mp z_{n+m-2}(\alpha) \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}$$

which also represents the central $100(1-\alpha)\%$ posterior credible interval of $\eta$ under the reference prior $\xi(\mu_1, \mu_2, \sigma^2) = \text{const.}/\sigma^2$. Suppose the observed data shows $n = 9, \bar{x} = -0.06, s_x = 0.92$ and $m = 11, \bar{y} = 1.84, s_y = 1.19$.

(a) What's the largest value of $\alpha$ that would produce an ML confidence interval with some support toward the theory $H_+$?

(b) How does the value of $\alpha$ from part (a) relate to the plausibility of $H_+$ under the reference-prior Bayesian analysis?