THE UNIVERSITY OF SYDNEY
SCHOOL OF MATHEMATICS AND STATISTICS

Solutions to Tutorial Week 13

STAT3023: Statistical Inference — Semester 2, 2023

Review exercises based on the geometric distribution

1. Recall that if a discrete random variable X has probability mass function (PMF) in the exponential family form

$$P_\theta(X = x) = e^{\theta t(x) - K(\theta) - M(x)} \tag{1}$$

then $E_\theta[t(X)] = K'(\theta)$ and $Var_\theta[t(X)] = K''(\theta)$. We call the parameter θ the "natural" or "canonical" parameter of the exponential family.

Suppose X has a geometric(p) distribution so that $P(X = x) = (1-p)^{x-1}p$ for $x = 1, 2, \dots$ By writing the PMF of X in exponential family form (1), deduce $E(X)$ and $\mathrm{Var}(X)$ as functions of p.

Solution: Since

$$P(X = x) = (1-p)^{x-1}p = e^{x\log(1-p)}\left(\frac{p}{1-p}\right)$$

we need to change parameters to $\theta = \log(1-p)$; equivalently $p = 1 - e^\theta$ and $1 - p = e^\theta$. Then the PMF takes the form

$$p_\theta(x) = e^{x\theta}\left(\frac{1-e^\theta}{e^\theta}\right) = e^{x\theta}\left(e^{-\theta} - 1\right) = e^{\theta x - \left[-\log\left(e^{-\theta}-1\right)\right]}$$

which is in form (1) with

$$K(\theta) = -\log\left(e^{-\theta} - 1\right).$$

Differentiating once gives

$$K'(\theta) = -\left\{\frac{-e^{-\theta}}{e^{-\theta}-1}\right\} = \frac{e^{-\theta}}{e^{-\theta}-1} = \frac{1}{1-e^\theta}$$

after multiplying top and bottom by e^θ. Differentiating again gives

$$K''(\theta) = -(1-e^\theta)^{-2}\left(-e^\theta\right) = \frac{e^\theta}{(1-e^\theta)^2}.$$

Expressing these in terms of p we get

$$E(X) = \frac{1}{p} \text{ and } Var(X) = \frac{1-p}{p^2}.$$

2. Suppose $X_1, \dots, X_n$ are iid geometric with $P(X_1 = x) = (1-p)^{x-1}p$ for $x = 1, 2, \dots$, but it is desired to estimate the *natural/canonical* parameter θ rather than p.

(a) Write down the likelihood $f_\theta(\mathbf{X})$ in terms of the natural parameter θ and hence obtain the score function $\ell'(\theta; \mathbf{X}) = \frac{\partial}{\partial\theta}\log f_\theta(\mathbf{X})$.

Solution:

$$f_\theta(\mathbf{X}) = \prod_{i=1}^n \left[e^{\theta X_i}(e^{-\theta} - 1)\right] = e^{\theta T}\left(e^{-\theta} - 1\right)^n$$

where $T = \sum_{i=1}^n X_i$. Thus the log-likelihood and its derivative with respect to θ are (respectively)

$$\ell(\theta; \mathbf{X}) = \log f_\theta(\mathbf{X}) = T\theta + n\log\left(e^{-\theta} - 1\right)$$

and

$$\ell'(\theta; \mathbf{X}) = T - \frac{ne^{-\theta}}{e^{-\theta}-1} = T - \frac{n}{1-e^\theta}.$$

(b) Determine the Cramér-Rao lower bound to the variance of an unbiased estimator of θ.

Solution:

$$\mathrm{Var}_\theta\left[\ell'(\theta;\mathbf{X})\right] = \mathrm{Var}_\theta(T) = n\mathrm{Var}_\theta(X_1) = \frac{ne^\theta}{(1-e^\theta)^2}\,.$$

Therefore for any unbiased estimator $\hat\theta$ of θ,

$$\mathrm{Var}_\theta\left(\hat\theta\right) \geq \frac{1}{\mathrm{Var}_\theta\left[\ell'(\theta;\mathbf{X})\right]} = \frac{(1-e^\theta)^2}{ne^\theta}\,.$$

(c) Derive the maximum-likelihood estimator $\hat\theta_{\mathrm{ML}}$ of θ.

Solution: Setting the score function equal to zero and solving gives

$$\begin{aligned} T &= \frac{n}{1-e^\theta} \\ e^\theta &= 1 - \frac{1}{\bar X} \\ \hat\theta_{\mathrm{ML}} &= \log\left(1 - \frac{1}{\bar X}\right)\,. \end{aligned}$$

3. Suppose $X_1, \ldots, X_{10}$ are iid geometric with $P(X = x) = (1-p)^{x-1}p$. Derive the UMP test at level 0.05 for testing $H_0\colon p = 0.5$ against the alternative $H_1\colon p > 0.5$. You may use the `R` output below and the fact that $T = \sum_{i=1}^n X_i$ has a *negative binomial* distribution and the CDF of $T - n$; specifically

$$P(T - n \leq x)$$

is given by the `R` function `pnbinom(x, n, p)`.

```
> x = 0:20
> cbind(x, pnbinom(x, 10, .5))
       x
 [1,]  0 0.0009765625
 [2,]  1 0.0058593750
 [3,]  2 0.0192871094
 [4,]  3 0.0461425781
 [5,]  4 0.0897827148
 [6,]  5 0.1508789063
 [7,]  6 0.2272491455
 [8,]  7 0.3145294189
 [9,]  8 0.4072647095
[10,]  9 0.5000000000
[11,] 10 0.5880985260
[12,] 11 0.6681880951
[13,] 12 0.7382664680
[14,] 13 0.7975635529
[15,] 14 0.8462718725
[16,] 15 0.8852385283
[17,] 16 0.9156812280
[18,] 17 0.9389609396
[19,] 18 0.9564207233
[20,] 19 0.9692858271
[21,] 20 0.9786130274
```

Solution: For any $p_0 < p_1$, the likelihood ratio statistic is

$$\frac{f_{p_1}(\mathbf{X})}{f_{p_0}(\mathbf{X})} = \prod_{i=1}^n \left[\frac{(1-p_1)^{X_i-1}p_1}{(1-p_0)^{X_i-1}p_0}\right] = \prod_{i=1}^n \left[\left(\frac{1-p_1}{1-p_0}\right)^{X_i}\left(\frac{p_1(1-p_0)}{(1-p_1)p_0}\right)\right] = \left(\frac{1-p_1}{1-p_0}\right)^T\left(\frac{p_1(1-p_0)}{(1-p_1)p_0}\right)^n\,.$$

Since $(1-p_1)/(1-p_0) < 1$, this is a *decreasing* function of T, thus it satisfies the monotone likelihood ratio property. The likelihood ratio test rejects for large values of the likelihood ratio statistic, which is equivalent to rejecting for *small* values of T. Specifically the test function is of the form

$$\delta(\mathbf{X}) = \begin{cases} 1 & \text{for } T < c \\ \gamma & \text{for } T = c \\ 0 & \text{for } T > c \end{cases}$$

where C and γ are chosen so that $E_{0.5}\left[\delta(X)\right] = 0.05$.

To determine the values C and γ, note that we must have

$$0.05 = E\left[\delta(\mathbf{X})\right] = P(T < C) + \gamma P(T = C) = (1-\gamma)P(T < c) + \gamma P(T \leq C)$$

thus

$$P(T < C) \leq 0.05 \leq P(T \leq C)\,.$$

From the `R` output we have

$$\begin{aligned} P(T - 10 \leq 3) &= P(T \leq 13) \approx 0.0461 \\ P(T - 10 \leq 4) &= P(T \leq 14) \approx 0.0898\,, \end{aligned}$$

so we must have $C = 14$. The value of γ is then given by

$$\gamma = \frac{0.05 - P(T < C)}{P(T = C)} \approx \frac{0.05 - 0.0461}{0.0898 - 0.0461} = \frac{0.0039}{0.0437} \approx 0.0892\,.$$

4. Suppose $X_1, \ldots, X_n$ are iid geometric with $P(X_1 = x) = (1-p)^{x-1}p$. Derive the Bayes estimator of p under squared-error loss using the $U[0,1]$ prior.

Solution: The product of the prior and likelihood is

$$w(p)f_p(\mathbf{X}) = (1-p)^{T-n}\,p^n = \text{const.}\ \frac{p^{(n+1)-1}(1-p)^{(T-n+1)-1}}{\text{beta}(n+1, T-n+1)}$$

so the posterior distribution is the $\text{beta}(n+1, T-n+1)$ distribution. Since we are using squared-error loss the Bayes estimator is the posterior mean which is

$$\frac{n+1}{(n+1)+(T-n+1)} = \frac{n+1}{T+2}\,.$$