Final Exam
Fall 2024

Stat 641 Probability Theory
Instructor: Garritt L. Page

Key

Show all work to receive credit. Credit (including partial credit) is given for work not final answers.

1. Prove the following. If X and Y are any two random variables, then $E(X) = E[E(X|Y)]$, provided the expectations exist (focus on the case of continuous random variables).

$$E(X) = \int_y \int_x x f_{XY}(x,y)\,dx\,dy$$

$$= \int_y \int_x x f_{X|Y}(x|y)\,dx\, f_Y(y)\,dy$$

$$= \int_y E(X|Y) f_Y(y)\,dy$$

$$= E_Y[E_{X|Y}(X|Y)]$$

2. Let the random variable $N \sim Poisson(\lambda)$ be the number of automobile accidents occurring on a given stretch of I-15 over a month. For each accident, let the random variable Y_i (for $i = 0, 1, 2, \ldots, N$) take the value 1 if the ith accident involved has at least one fatality, and let Y_i take the value 0 otherwise (it is a Bernoulli trial). Let $P(Y_i = 1) = \theta$ (where $0 < \theta < 1$) and further assume that all Y_i are mutually independent.

Suppose we are interested in the total number of automobile accidents involving fatalities on that stretch of I-15 during the specified time period. Given $N = n$, we define $T = \sum_{i=1}^{n} Y_i$. Note that T is dependent on N.

(a) What is the distribution of $T|N = n$? Justify your answer.

Note that $Y_i \sim Ber(\theta)$. $M_{Y_i}(t) = (1-\theta) + \theta e^t$ $i = 1, \ldots, N$

$T = \sum_{i=1}^{N} Y_i \Rightarrow T \mid N = n = \sum_{i=1}^{n} Y_i$

$M_T(t) = E(e^{tT}) = E(e^{t\sum Y_i}) = E\left(\prod_{i=1}^{n} e^{tY_i}\right) = \prod_{i=1}^{n} E(e^{tY_i}) = [E(e^{tY_i})]^n = [M_Y(t)]^n$

$M_T(t) = [M_Y(t)]^n = ((1-\theta) \quad \theta e^t)^n$

This is the MGF for a $Bin(n, \theta)$ so $T \mid N = n \sim Bin(n, \theta)$

(b) What is $E(T)$?

$E(T) = E(E(T \mid N)) = E(N\theta) = \theta E(N) = \theta\lambda$

(c) What is $Var(T)$?

$$Var(T) = E(Var(T|N)) + Var(E(T|N))$$
$$= E(N\theta(1-\theta)) + Var(N\theta)$$
$$= \lambda\theta(1-\theta) + \theta^2\lambda$$
$$= \lambda\theta - \lambda\theta^2 + \theta^2\lambda$$
$$= \lambda\theta$$

(d) What is $Cov(N, T)$?

$$E(NT) - E(N)E(T)$$
$$= E(N\,E(T|N)) - \lambda(\theta\lambda)$$
$$= E(N N\theta) - \lambda^2\theta$$
$$= \theta E(N^2) - \lambda^2\theta$$
$$= \theta(Var(N) + E(N)^2) - \lambda^2\theta$$
$$= \theta(\lambda + \lambda^2) - \lambda^2\theta$$
$$= \theta\lambda$$

(e) What is $P(T = 0)$, the probability there are no fatal accidents on that stretch of road during the time period?

$$P(T=0) = \sum_{n=0}^{\infty} P(T=0, N=n)$$

$$= \sum_{n=0}^{\infty} P(T=0 \mid N=n)\, P(N=n)$$

$$= \sum_{n=0}^{\infty} \binom{n}{0} \theta^0 (1-\theta)^{n-0} \frac{e^{-\lambda} \lambda^n}{n!}$$

$$= \sum_{n=0}^{\infty} \frac{[(1-\theta)\lambda]^n e^{-\lambda}}{n!}$$

$$= \sum_{n=0}^{\infty} \frac{[(1-\theta)\lambda]^n e^{-\lambda}}{n!} e^{\lambda\theta} e^{-\lambda\theta}$$

$$= e^{-\lambda\theta} \sum_{n=0}^{\infty} \frac{[(1-\theta)\lambda]^n e^{-\lambda(1-\theta)}}{n!}$$

$$= e^{-\lambda\theta}$$

3. Let $X_1, \ldots, X_n$ be a random sample from a $UN(0,1)$ distribution. What is the pdf of the sample range $R = X_{(n)} - X_{(1)}$?

$$f_{X_{(1)}, X_{(n)}}(x_1, x_n) = n(n-1)(x_n - x_1)^{n-2} \qquad 0 < x_1 < x_n < 1$$

Transform to $(R, Y = X_{(n)})$ $\qquad R = X_{(n)} - X_{(1)} \qquad Y = X_{(n)}$

$\Rightarrow X_{(n)} = Y$ and $X_{(1)} = Y - R$

$$\frac{dX_{(n)}}{dY} = 1 \qquad \frac{dX_{(n)}}{dR} = 0$$

$$\frac{dX_{(1)}}{dY} = 1 \qquad \frac{dX_{(1)}}{dR} = -1$$

$$f_{RY}(r, y) = n(n-1)(y - y - r)^{n-2} \quad \text{for } Y \in (0,1) \text{ and } 0 < R < Y < 1$$

$$= n(n-1)(r)^{n-2}$$

$$f_R(r) = \int_R^1 n(n-1)(r)^{n-2}\, dy$$

$$= n(n-1) r^{n-2} y \Big|_r^1$$

$$= n(n-1) r^{n-2}(1-r) \qquad r \in (0,1)$$

$$\boxed{= n(n-1)[r^{n-2} - r^{n-1}]}$$

$$\int_0^1 n(n-1)[r^{n-2} - r^{n-1}]\, dr$$

$$= n(n-1) \left. \frac{1}{n-1} r^{n-1} \right|_0^1 - \left. \frac{1}{n} r^n \right|_0^1$$

$$= n(n-1)\left[\frac{1}{n-1} - \frac{1}{n}\right]$$

$$= n(n-1)\left[\frac{n-(n-1)}{n(n-1)}\right] = 1$$

4. The Inverse Gaussian distribution is a two-parameter continuous distribution with pdf

$$f(x|\mu,\lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left\{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}.\right\} \quad x>0;\ \mu>0;\ \lambda>0.$$

(a) Show that this family can be written in exponential family form. Identify each of the h, c, t, and w functions. Assume both μ and λ are unknown. Please put constants in the functions of parameters rather than in the functions of x.

$$f(x|\mu,\lambda) = \left(\frac{1}{x^3}\right)^{\frac{1}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{\lambda}{2\mu^2 x}(x^2 - 2\mu x + \mu^2)\right\}$$

$$= \left(\frac{1}{x^3}\right)^{\frac{1}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{\lambda}{2\mu^2}x + \frac{\lambda}{\mu} - \frac{\lambda}{2}\frac{1}{x}\right\}$$

$$= \left(\frac{1}{x^3}\right)^{\frac{1}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\frac{\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda}{2\mu^2}x - \frac{\lambda}{2}\frac{1}{x}\right\}$$

$$h(x) = \left(\frac{1}{x^3}\right)^{\frac{1}{2}} I[x>0]$$

$$c(\mu,\lambda) = \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\frac{\lambda}{\mu}\right\}$$

$$t_1(x) = x \qquad t_2(x) = \frac{1}{x}$$

$$w_1(\mu,\lambda) = -\frac{\lambda}{2\mu^2} \qquad w_2(\mu,\lambda) = -\frac{\lambda}{2}$$

(b) Using part (a), find $E(1/X)$.

This can be found using the cannonical form

$$\eta_1 = w_1(\mu,\lambda) = -\frac{\lambda}{2\mu^2} \qquad \eta_2 = w_2(\mu,\lambda) = -\frac{\lambda}{2}$$

$$\Rightarrow \lambda = -2\eta_2 \quad \text{and} \quad \mu^2 = \frac{\eta_2}{\eta_1} \qquad (\text{note } \eta_1 < 0 \text{ and } \eta_2 < 0)$$

$$c^*(\eta_1,\eta_2) = \left(\frac{-2\eta_2}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\frac{-2\eta_2}{\sqrt{\frac{\eta_2}{\eta_1}}}\right\}$$

$$= \left(-\frac{\eta_2}{\pi}\right)^{\frac{1}{2}} \exp\left\{-(\eta_2\eta_1)^{\frac{1}{2}}\right\}$$

$$E\left(\frac{1}{X}\right) = -\frac{\partial}{\partial\eta_2}\log\left[\left(-\frac{\eta_2}{\pi}\right)^{\frac{1}{2}} \exp\left\{-(\eta_2\eta_1)^{\frac{1}{2}}\right\}\right]$$

$$= -\frac{\partial}{\partial\eta_2}\left\{\frac{1}{2}\log(-\eta_2) - \frac{1}{2}\log(\pi) - (\eta_2\eta_1)^{\frac{1}{2}}\right\}$$

$$= -\left[\frac{1}{2}\left(-\frac{1}{\eta_2}\right)(-1) - \frac{1}{2}(\eta_2\eta_1)^{-\frac{1}{2}}\eta_1\right]$$

$$= \frac{1}{2}\left[-\frac{1}{\eta_2} + \sqrt{\frac{\eta_1}{\eta_2}}\right] = \frac{1}{2}\left[-\frac{1}{-\frac{\lambda}{2}} + \sqrt{\frac{-\frac{\lambda}{2\mu^2}}{-\frac{\lambda}{2}}}\right] = \frac{1}{2}\left[\frac{2}{\lambda} + \frac{2}{\mu}\right] = \frac{1}{\lambda} + \frac{1}{\mu}$$

5. Let X_i, $i = 1, 2, \ldots$, be *iid* $Ber(p)$ random variables and let $Y_n = \frac{1}{n}\sum_{i=1} X_i$.

(a) Show that $\log\left(\frac{Y_n}{1-Y_n}\right) \xrightarrow{p} \log\left(\frac{p}{1-p}\right)$

$$E(Y_n) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = \frac{1}{n} np = p$$

$$Var(Y_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

Let $\varepsilon > 0$ then $Pr(|Y_n - p| \geq \varepsilon) = Pr((Y_n - p)^2 \geq \varepsilon^2) \leq \frac{Var(Y_n)}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}$

$$\lim_{n\to\infty} P(|Y_n - p| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{p(1-p)}{n\varepsilon^2} = 0$$

$\Rightarrow \quad Y_n \xrightarrow{p} p$

by the cont. mapping theorem

$$\log\left(\frac{Y_n}{1-Y_n}\right) \xrightarrow{p} \log\left(\frac{p}{1-p}\right)$$

(b) What is the limiting distribution of $\sqrt{n}\left[\log\left(\frac{Y_n}{1-Y_n}\right) - \log\left(\frac{p}{1-p}\right)\right]$? Explain.

by CLT $\sqrt{n}(Y_n - p) \xrightarrow{d} N(0, p(1-p))$

by delta method $\sqrt{n}(g(Y_n) - g(p)) \xrightarrow{d} N(0, (g'(p))^2 p(1-p))$

Let $g(Y_n) = \log\left(\frac{Y_n}{1-Y_n}\right)$ and $g(p) = \log\left(\frac{p}{1-p}\right)$

Then $g'(p) = \frac{1}{\frac{p}{1-p}}\left[p(-1)(1-p)^{-2}(-1) + (1-p)^{-1}(1)\right]$

$$= \frac{1-p}{p}\left[\frac{p}{(1-p)^2} + \frac{1}{1-p}\right]$$

$$= \frac{1}{1-p} + \frac{1}{p} = \frac{p + (1-p)}{p(1-p)} = \frac{1}{p(1-p)}$$

thus $\sqrt{n}\left(\log\left(\frac{Y_n}{1-Y_n}\right) - \log\left(\frac{p}{1-p}\right)\right) \xrightarrow{d} N\left(0, \left(\frac{1}{p(1-p)}\right)^2 \frac{p(1-p)}{1}\right)$

$$= N\left(0, \frac{1}{p(1-p)}\right)$$

6. Assume that $X|Y \sim Poisson(Y)$, where $Y \sim Gamma(\alpha, \beta)$. What is $M_X(t)$ (i.e., the marginal moment generating function of X)? Bonus points if you can name what distribution it is.

$X \sim NB$

$$M_X(t) = E(e^{tX}) = E[E(e^{tX}|Y)]$$

$$= E[e^{Y(e^t-1)}]$$

$$= \int_0^\infty e^{Y(e^t-1)} f_Y(y)\,dy$$

$$= \int_0^\infty e^{Y(e^t-1)} \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-\frac{1}{\beta}y}\,dy$$

$$= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y(\frac{1}{\beta}-e^t+1)}\,dy$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \frac{\Gamma(\alpha)}{(\frac{1}{\beta}-e^t+1)^\alpha} \int_0^\infty \frac{(\frac{1}{\beta}-e^t+1)^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-y(\frac{1}{\beta}-e^t+1)}\,dy$$

$$= \frac{1}{[\beta(\frac{1}{\beta}-e^t+1)]^\alpha}$$

$$= \left(\frac{1}{1-\beta e^t+\beta}\right)^\alpha$$

$$= \left(\frac{1}{1+\beta-\beta e^t}\right)^\alpha$$

Let $\beta = \frac{1-p}{p}$ then $M_X(t) = \left(\frac{1}{1+\frac{1-p}{p}-\frac{1-p}{p}e^t}\right)^\alpha$

$$= \left(\frac{p}{p+1-p-(1-p)e^t}\right)^\alpha$$

$$= \left(\frac{p}{1-(1-p)e^t}\right)^\alpha$$

which is the MGF of a negative binomial with $p = \frac{1}{\beta+1}$

7. Suppose that $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ for $i = 1, \ldots, n$ and $Z_i \overset{iid}{\sim} N(0, 1)$ for $i = 1, \ldots, k$, and all variables are independent. For (a) - (d) state the distribution if it is a "named" distribution. Otherwise simply put "unknown". "Known" here means that it is a distribution that has been covered in class or is on the Casella & Berger list provided. Parameter values must also be provided. Make sure to defend your answers.

(a) $\dfrac{Z_1}{Z_2}$ → Cauchy see dist. sheet

(b) $\dfrac{\bar{X}}{\bar{Z}}$ unknown

(c) $k\bar{Z}^2$ Note $k\bar{Z}^2 = (\sqrt{k}\,\bar{Z})^2 = Z^2 \sim \chi^2_{(1)}$

$\bar{Z} \sim N(0, \frac{1}{k}) \Rightarrow \sqrt{k}\,\bar{Z} \sim N(0, 1)$

(d) $\frac{1}{\sigma^2}\bar{X} + \frac{1}{k}\sum_{i=1}^{k} Z_i$

Note $\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \frac{1}{\sigma^2}\bar{X} \sim N(\frac{\mu}{\sigma^2} + \frac{1}{\sigma^2 n})$

$\frac{1}{k}\sum_{i=1}^{n} Z_i = \bar{Z} \sim N(0, \frac{1}{k})$

since $\bar{X}, \bar{Z}$ are ind

$\frac{1}{\sigma^2}\bar{X} + \bar{Z} \sim N\left(\frac{\mu}{\sigma^2} + 0, \frac{1}{\sigma^2 n} + \frac{1}{k}\right)$