Stat 542-3 (1) - Ph.D. Prelim Exam - Spring 1998

For a discrete random variable X taking values on the nonnegative integers the generating function is defined as $h_X(s) = E(s^X) = \sum_{j=0}^{\infty} \Pr(X = j)s^j$ for those values s for which the sum converges.

(a) Prove that $h_X(s)$ converges for $-1 \leq s \leq 1$.

(b) Show that $h'_X(1) = \lim_{s \uparrow 1} h'_X(s)$ is equal to the mean of the random variable X.

(c) Obtain an expression for the variance of X (assuming the variance is finite) in terms of $h_X(s)$ or its derivatives.

(d) If X and Y are independent random variables with generating functions $h_X(s)$ and $h_Y(s)$ then find the generating function of $X + Y$.

(e) Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having generating function $f(s)$. Let $S_N = X_1 + X_2 + \cdots + X_N$ where N is a nonnegative integer-valued random variable independent of the X_i's. We take $S_N = 0$ if $N = 0$. Let the generating function of N be $g(s)$.

   i. Find the generating function of S_N conditional on $N = n$.

   ii. Show that the generating function of S_N is $g(f(s))$.

(f) Show that the generating function for a Poisson random variable with mean λ is $e^{\lambda(s-1)}$.

(g) In the setup of part (e), suppose that the X_i's are i.i.d. Bernoulli trials with probability of success p and N is a Poisson random variable with mean λ.

   i. Find the distributions of S_N and $N - S_N$, and their means and variances.

   ii. Show that S_N and $N - S_N$ are independent by computing $\Pr(S_N = u, N - S_N = v)$.

(h) In the setup of part (e), suppose that the X_i's are i.i.d. Poisson random variables with mean μ and N is a Poisson random variable with mean λ. Find the mean and variance of the distribution of S_N.

Stat 542-3 (1) Solution - Ph.D. Prelim Exam - Spring 1998

ⓐ $h_X(1) = \sum_{j=0}^{\infty} Pr(X=j) = 1$ so $h_X(s)$ is convergent at $s=1$.
For $-1 \le s < 1$ we have $|Pr(X=j)\, s^j| \le Pr(X=j)$ so $h_X(s)$ converges for s in this interval.

ⓑ $h_X'(s) = \sum_{j=1}^{\infty} j\, Pr(X=j)\, s^{j-1}$. If $EX < \infty$ then $\lim_{s \uparrow 1} h_X'(s) = h_X'(1) = \sum_j j\, Pr(X=j)$
$= EX$

ⓒ $h_X''(1) = \sum_{j=2}^{\infty} j(j-1)\, Pr(X=j) = E[X(X-1)]$
Then $\operatorname{Var} X = EX^2 - (EX)^2 = h_X''(1) + h_X'(1) - h_X'(1)^2$

ⓓ $h_{X+Y}(s) = E(s^{X+Y}) = E(s^X)E(s^Y) = h_X(s)\, h_Y(s)$

ⓔ i. $E[S^{X_1 + \cdots + X_N} \mid N=n] = [f(s)]^n$ (apply result from ⓓ)
ii. $h_{S_N}(s) = \sum_{n=0}^{\infty} E[S^{X_1+\cdots+X_N} \mid N=n]\, Pr(N=n) = \sum_{n=0}^{\infty} Pr(N=n)\,[f(s)]^n = g(f(s))$

ⓕ $X \sim Poi(\lambda)$ $h_X(s) = \sum_{j=0}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!} s^j = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}$

ⓖ i. $X \sim Bern(p)$ $h_X(s) = (1-p) + ps$
From ⓔ + ⓕ we find $h_{S_N}(s) = e^{\lambda(h_X(s)-1)} = e^{\lambda p(s-1)}$
$\Rightarrow$ $S_N \sim Poi(\lambda p)$ $ES_N = \operatorname{Var} S_N = \lambda p$
If we apply the same argument to $Y = 1-X$ with $N - S_N = Y_1 + \cdots + Y_N$ we find $N - S_N \sim Poi(\lambda(1-p))$ with $E(N-S_N) = \operatorname{Var}(N-S_N) = \lambda(1-p)$

ii. $Pr(S_N = u, N - S_N = v) = Pr(S_N = u, N = u+v)$
$= Pr(N = u+v)\, Pr(S_N = u \mid N = u+v)$
$= \frac{\lambda^{u+v} e^{-\lambda}}{(u+v)!} \frac{(u+v)!}{u!\,v!} p^u (1-p)^v = \frac{e^{-\lambda p}(\lambda p)^u}{u!} \frac{e^{-\lambda(1-p)}(\lambda(1-p))^v}{v!}$
$= Pr(S_N = u)\, Pr(N - S_N = v)$ so S_N and $N - S_N$ are independent.

ⓗ From ⓔ + ⓕ $h_{S_N}(s) = e^{\lambda(h_X(s)-1)} = e^{\lambda(e^{\mu(s-1)}-1)}$
The mean and variance of S_N can be found from ⓑ + ⓒ OR
$E(S_N) = E(E(S_N \mid N)) = E[N\mu] = \lambda\mu$
$\operatorname{Var}(S_N) = E(\operatorname{Var}(S_N \mid N)) + \operatorname{Var}(E(S_N \mid N)) = E[N\mu] + \operatorname{Var}[N\mu] = \lambda\mu + \lambda\mu^2$

Stat 542-3 (2) - Ph.D. Prelim Exam - Spring 1998

Assume that X has the exponential distribution, i.e., its pdf is $f_X(x) = \theta e^{-\theta x}$ for $0 < x < \infty$ and 0 otherwise, where $\theta > 0$.

(a) Derive the mean and variance of X.

(b) Suppose that we have an i.i.d. sample $X_1, \dots, X_n$ of observations ($n \geq 2$) from the exponential distribution described at the top of the page.

i. Find the maximum likelihood estimator $\hat{\theta}$ for θ.

ii. Show that the distribution of $\sum_{i=1}^{n} X_i$ is a gamma distribution. (Recall if $Y \sim \text{Gamma}(\alpha, \beta)$ then its p.d.f. is $f_Y(y) = \beta^\alpha y^{\alpha-1} e^{-\beta y}/\Gamma(\alpha)$ for $0 < y < \infty$ and 0 otherwise).

iii. Find the expected value of the maximum likelihood estimator for θ. Is the maximum likelihood estimator for θ unbiased?

iv. Find an exact 95% confidence interval for θ based on $X_1, \dots, X_n$.

v. Identify the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta)$.

(c) Let $[x]$ denote the largest integer less than or equal to x. Define $Y = [X]$ where X has the exponential distribution described at the top of the page. Show that Y has a geometric distribution and identify the mean and variance of Y.

(d) Suppose that in part (b) we are not able to observe the actual values $X_1, \dots, X_n$, but instead observe the values $Y_1, \dots, Y_n$ with $Y_i = [X_i]$.

i. Find the maximum likelihood estimator for θ.

ii. Find the distribution of $\sum_{i=1}^{n} Y_i$.

(e) With the setup as in (d), intuition should suggest that we get more information about θ when we observe the X_i's than when we observe the Y_i's. Check this intuition by considering the Fisher information in each case.

Stat 542-3 (2a) Solution - Ph.D. Prelim Exam - Spring 1998

ⓐ $EX = \int_0^\infty \theta x e^{-\theta x}\,dx = 1/\theta$

$EX^2 = \int_0^\infty \theta x^2 e^{-\theta x}\,dx = 2/\theta^2 \Rightarrow \quad Var X = 1/\theta^2$

ⓑ $f(x_1, \dots x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) = \theta^n e^{-\theta \Sigma x_i} = L(\theta)$

i. $\frac{d \log L}{d\theta} = \frac{n}{\theta} - \Sigma X_i = 0 \Rightarrow \quad \hat{\theta} = n/\Sigma X_i = 1/\bar{x}$

ii. Using MGFs $\quad M_{X_i}(t) = E(e^{tX}) = \theta/(\theta - t)$

$M_{\Sigma X_i}(t) = (\theta/(\theta-t))^n$ which is the MGF for a Gamma(n,θ) dist^n.

iii. $E(\hat{\theta}) = E\left(\frac{n}{\Sigma X}\right) = \int_0^\infty \frac{n}{y} \frac{\theta^n y^{n-1} e^{-\theta y}}{\Gamma(n)}\,dy = \frac{n\theta}{n-1}$, $\quad \hat{\theta}$ is not unbiased

iv. If $\Sigma X_i \sim$ Gamma(n,θ) then $\theta \Sigma X_i \sim$ Gamma$(n,1)$

Let $\Gamma_{\alpha,\beta,p} = p^{th}$ percentile of the Gamma(α,β) distribution.

Then $\left(\frac{\Gamma_{n,1,.025}}{\Sigma X_i}, \frac{\Gamma_{n,1,.975}}{\Sigma X_i} \right)$ is a 95% CI for θ

v. MLE $\hat{\theta} \approx N(\theta, I^{-1})$ where $I = E\left[-\frac{d^2 \log L}{d\theta^2}\right] = n/\theta^2$

$\Rightarrow \quad \sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \theta^2)$

ⓒ $P(Y=y) = P(y \le X \le y+1) = \int_y^{y+1} \theta e^{-\theta x}\,dx = e^{-\theta y}(1 - e^{-\theta})$

If $X \sim$ Geom(p) $\quad f_X(x) = q^x p \quad \begin{matrix} x = 0, 1, \dots \\ q = 1-p \end{matrix}$ then $EX = q/p \quad Var X = q/p^2$

Therefore $Y \sim$ Geom$(1 - e^{-\theta})$ with $EY = 1/(e^\theta - 1) \quad Var\,Y = e^\theta/(e^\theta - 1)^2$

ⓓ i. $f(y_1, \dots y_n \mid \theta) = \prod_{i=1}^{n} e^{-\theta y_i}(1 - e^{-\theta}) = e^{-\theta \Sigma y_i}(1 - e^{-\theta})^n = L(\theta)$

$\frac{d \log L}{d\theta} = -\Sigma Y_i + n e^{-\theta}/(1 - e^{-\theta}) = 0 \Rightarrow \quad \hat{\theta} = \begin{cases} \ln\left(1 + \frac{1}{\bar{y}}\right) & \text{if } \Sigma Y_i > 0 \\ \infty & \text{if } \Sigma Y_i = 0 \end{cases}$

ii. If $Y_i \sim$ Geom(p) then $Z = \Sigma Y_i \sim$ NegBin(n,p), $P(Z=z) = \binom{z+n-1}{z} e^{-\theta z}(1 - e^{-\theta})^n$

ⓔ Note that more information (in the Fisher sense) implies a lower asymptotic variance.

X: $\quad I_X = E\left[-\frac{d^2 \log f(x_1 \dots x_n \mid \theta)}{d\theta^2}\right] = n/\theta^2$

Y: $\quad I_Y = E\left[-\frac{d^2 \log f(y_1 \dots y_n \mid \theta)}{d\theta^2}\right] = ne^\theta/(e^\theta - 1)^2 = n/(e^{\theta/2} - e^{-\theta/2})^2$

$= n/\left(\theta + \frac{2}{3!}\left(\frac{\theta}{2}\right)^3 + \frac{2}{5!}\left(\frac{\theta}{2}\right)^5 + \dots\right) < n/\theta^2 = I_X$

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with common probability density function (pdf) $f_\theta(x)$, $\theta > 0$, where

$$f_\theta(x) = \begin{cases} \frac{3x^2}{\theta^3} & \text{if} \quad 0 < x < \theta, \\ 0 & \text{otherwise.} \end{cases}$$

Also, suppose that the parameter θ is random with prior pdf

$$g(\theta) = \begin{cases} e^{-\theta}\theta^{3n}/\{(3n)!\} & \text{if} \quad \theta > 0, \\ 0 & \text{otherwise.} \end{cases}$$

i. Write down the joint pdf of $X_1, \ldots, X_n$ and θ.

ii. Find the marginal distribution of $X_1, \ldots, X_n$.

iii. Show that the posterior density function (i.e., the conditional pdf) of θ given $X_i = x_i,\ i = 1, \ldots, n$, is as follows:

$$g_{\theta|\mathbf{x}}(\theta) = \begin{cases} \exp(-(\theta - x_{(n)})) & \text{if} \quad \theta > x_{(n)}, \\ 0 & \text{otherwise,} \end{cases}$$

where $x_{(n)} = \max_{1 \le i \le n} x_i$ and $\mathbf{x} = (x_1, \ldots, x_n)'$.

iv. Find the Bayes' estimator of θ under the squared error loss function.

v. Find a $100(1-\alpha)\%$ highest posterior density (HPD) credible set Λ_α for θ, where $0 < \alpha < 1$.
(Recall that Λ_α is a $100(1-\alpha)\%$ HPD credible set for θ if $\Lambda_\alpha = \{\theta : g_{\theta|\mathbf{x}}(\theta) > c\}$ for some real number c and $P(\theta \in \Lambda_\alpha | X_i = x_i,\ i = 1,, \ldots, n) = 1 - \alpha$.)

vi. Find $P_\theta(\theta \in \Lambda_\alpha)$.

vii. Is it true that the HPD credible set Λ_α can be regarded as a $100(1-\alpha)\%$ confidence region for θ ? Briefly justify your answer.

(i)

$$f(\underset{\sim}{x}, \theta) = 3^n \theta^{-3n} \prod_{i=1}^{n} \left\{ x_i^2 \, I(0 < x_i < \theta) \right\} \frac{e^{-\theta} \theta^{3n}}{(3n)!} \cdot I(\theta >$$

$$= \frac{3^n}{3n!} \left(\prod_{i=1}^{n} x_i^2 \right) e^{-\theta} \cdot I(0 < x_{(1)} < x_{(n)} < \theta$$

where $x_{(1)} = \min_{1 \le i \le n} x_i$ and $x_{(n)} = \max_{1 \le i \le n} x_i$.

(ii) The marginal pdf of $X_1, \cdots, X_n$ is given by

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = \int_{-\infty}^{\infty} f(\underset{\sim}{x}, \theta) \, d\theta$$

$$= \left\{ \int_{x_{(n)}}^{\infty} \frac{3^n}{3n!} \left(\prod_{i=1}^{n} x_i^2 \right) e^{-\theta} \, d\theta \right\} \cdot I(0 < x$$

$$= \frac{3^n}{3n!} \left(\prod_{1}^{n} x_i^2 \right) \cdot e^{-x_{(n)}} \cdot I(x_{(1)} > 0)$$

(iii) For $x_{(1)} > 0$,

$$g_{\theta|\underset{\sim}{x}}(\theta) = f(\underset{\sim}{x}, \theta) / f_x(\underset{\sim}{x})$$

$$= \frac{\frac{3^n}{3n!}\left(\Pi x_i^2\right) . e^{-\theta} \quad I(x_{(n)} < \theta)}{\frac{3^n}{3n!}\left(\Pi x_i^2\right) e^{-x_{(n)}}}$$

$$= e^{-(\theta - x_{(n)})} \quad I(\theta > x_{(n)})$$

(iv) The Bayes' estimator of θ is given by

$$E_{\theta|\underset{\sim}{x}}(\theta) = \int \theta \; g_{\theta|\underset{\sim}{x}}(\theta) \, d\theta$$

$$= \int_{X_{(n)}}^{\infty} \theta \, e^{-(\theta - X_{(n)})} \, d\theta$$

$$= X_{(n)} + 1$$

(v)

Note that a $(1-\alpha)$ HPD region is given by $\Lambda_\alpha \equiv \{ \theta \in (0,\infty) : g_{\theta|\underset{\sim}{x}}(\theta) > c \}$ where the constant $c > 0$ is such that

$$P_{\theta|\underset{\sim}{x}} \left(\theta \in \Lambda_\alpha \right) = 1-\alpha$$

Since $g_{\theta|\underset{\sim}{x}}(\theta)$ is a decreasing function over $(X_{(n)}, \infty)$, $g_{\theta|\underset{\sim}{x}}(\theta) > c \iff X_{(n)} < \theta < K$ (where $K = K(c)$ is defined by the equation $e^{-(K - X_{(n)})} = c$). Hence, a $(1-\alpha)$ HPD region for θ is given by $\Lambda_\alpha = [X_{(n)}, K]$ where

$$P_{\theta|\underset{\sim}{x}} \left(\theta < K \right) = 1-\alpha \iff \int_{X_{(n)}}^{K} e^{-(\theta - X_{(n)})} d\theta = 1-\alpha$$

$$\iff 1 - e^{-(K - X_{(n)})} = 1-\alpha \iff K = X_{(n)} - \log\alpha .$$

(vi)

$$P_\theta(\theta \in \Lambda_\alpha)$$

$$= P_\theta(\ X_{(n)} \le \theta \le X_{(n)} - \log\alpha)$$

$$= P_\theta(\ \theta + \log\alpha \le X_{(n)} \le \theta)$$

$$= 1 - P_\theta(\ X_{(n)} < \theta + \log\alpha)$$

$$= 1 - \left[P_\theta(X_1 < \theta + \log\alpha)\right]^n$$

$$= 1 - \left[\frac{3}{\theta^3} \cdot \int_0^{(\theta+\log\alpha)\vee 0} x^2\, dx\right]^n$$

$$= 1 - \left[\frac{3}{\theta^3} \cdot \left((\theta+\log\alpha)\vee 0\right)^3/3\right]^n$$

$$= 1 - \left[0 \vee (1 + \theta^{-1}\log\alpha)\right]^{3n}$$

NOTE: for $0 < \theta \le -\log\alpha$, $P_\theta(\theta \in \Lambda_\alpha) = 1$, and for $\theta > -\log\alpha$,

$1 + \theta^{-1}\log\alpha \in (0,1) \Rightarrow P_\theta(\theta \in \Lambda_\alpha) \longrightarrow 1$ as $n \to \infty$. Thus,

(vii) a $(1-\alpha)$ HPD region can not be interpreted as a $(1-\alpha)$ confidence region for θ.