Stat 641 Fall 2021

Solutions for Assignment 2

(1.) (10 points) In the following expressions let m be a non-negative integer. Using the expression

$\sum_{k=0}^{m} ab^k = a\frac{1-b^{m+1}}{1-b}$, we have with a = p, b = 1- p, [y] = greatest integer $\leq$ y

(a.) For $y < 0$, F(y) = 0; for $y \geq 0$,

$$F(y) = P[Y \leq y] = \sum_{k=0}^{[y]} p(1-p)^k = 1-(1-p)^{[y]+1} = \begin{cases} 0 & \text{if } y < 0 \\ p & \text{for } 0 \leq y < 1 \\ 1-(1-p)^2 & \text{for } 1 \leq y < 2 \\ 1-(1-p)^3 & \text{for } 2 \leq y < 3 \\ 1-(1-p)^4 & \text{for } 3 \leq y < 4 \\ \vdots & \end{cases}$$

(b.)

$$\text{Using the definition of } Q(u),\ Q(u) = inf(y : F(y) \geq u) \ \Rightarrow$$

$$Q(u) = \text{ smallest nonnegative integer } y_u \text{ such that } 1-(1-p)^{y_u+1} \geq u \text{ with } Q(0) = 0$$

$$Q(u) = \text{ smallest nonnegative integer } y_u \text{ such that } y_u \geq \frac{log(1-u)}{log(1-p)} - 1 \text{ with } Q(0) = 0$$

That is,

$$Q(u) = \begin{cases} 0 & \text{if } u \leq 0 \\ 0 & \text{for } 0 < u \leq p \\ 1 & \text{for } p < u \leq 1-(1-p)^2 \\ 2 & \text{for } 1-(1-p)^2 < u \leq 1-(1-p)^3 \\ 3 & \text{for } 1-(1-p)^3 < u \leq 1-(1-p)^4 \\ \vdots & \end{cases}$$

Using the above expression with p = .4, Q(.8) = 3

(2.) (20 Points) (a.) Let $W = \frac{Y-\theta}{\alpha}$ then the pdf of W is

$$f_W(w) = \alpha f(\theta + \alpha w) = \alpha \frac{\gamma}{\alpha^\gamma}((\theta + \alpha w) - \theta)^{\gamma-1} e^{-\left(\frac{(\theta+\alpha w)-\theta}{\alpha}\right)^\gamma} \text{ for } \theta + \alpha w \geq \theta \Rightarrow$$

$$f_W(w) = \begin{cases} \gamma w^{\gamma-1} e^{-w^\gamma} & \text{for } w \geq 0 \\ 0 & \text{for } w < 0 \end{cases}$$

Because the expression for $f_W(w)$ does not contain $(\theta,\ \alpha)$, we can conclude that $(\theta,\ \alpha)$ are location-scale parameters for the family of distributions.

(b.) To find the quantile function, set

$$u = F(y_u) = 1 - e^{-\left(\frac{y_u-\theta}{\alpha}\right)^\gamma}$$

and solve for y_u. In this case,

$$y_u = \theta + \alpha(-log(1-u))^{1/\gamma} \quad \Rightarrow \quad Q(u) = \theta + \alpha(-log(1-u))^{1/\gamma}$$

(c.) With $\theta = 5,\ \gamma = 4,\ \alpha = 20,\quad P(Y > 23) = 1 - P(Y \leq 23) = 1 - F(23) = e^{-\left(\frac{23-5}{20}\right)^4} = .5189$

(d.) With $\theta = 5,\ \gamma = 4,\ \alpha = 20,\quad Q(.25) = 5 + 20(-log(1-.25))^{1/4} = 19.647$

(3.) (10 points) (a.) Let $W = Y/\beta$. The pdf of W is

$$f_W(w) = \beta f(\beta w) = \beta \frac{\gamma}{\beta}(\beta w)^{\gamma-1} e^{-(\beta w)^\gamma/\beta} = \gamma \beta^{\gamma-1} w^{\gamma-1} e^{-\beta^{\gamma-1} w^\gamma} \quad \text{for} \quad w > 0$$

Because the expression for the pdf of W contains β, β cannot be a scale parameter for the given family of distributions.

(b.) With $W = Y/\alpha$, the pdf of W is given by

$$f_W(w) = \alpha f(\alpha w) = \alpha \frac{\gamma}{\alpha}(\alpha w/\alpha)^{\gamma-1} e^{-(\alpha w/\alpha)^\gamma} = \gamma w^{\gamma-1} e^{-w^\gamma} \quad \text{for} \quad w > 0$$

The expression for f_W is free of α, therefore α is a scale parameter.

(4.) (10 points)

(a.) Let X be the number of emissions in a week. X has a Poisson distribution with $\lambda = 0.15$.

$$P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \frac{e^{-0.15}(0.15)^0}{0!} - \frac{e^{-0.15}(0.15)^1}{1!} = 1 - .8607 - .1291 = 0.0102$$

Using the R-function dpois, $P(X \geq 2) = 1 - dpois(0, .15) - dpois(1, .15) = 1 - .8607 - .1291 = 0.0102$

(b) Let Y be the number of emissions in a year. Y has a Poisson distribution with $\lambda = 0.15 \times 52 = 7.8$.

$$P(Y \geq 2) = 1 - P(Y=0) - P(Y=1) = 1 - \frac{e^{-7.8}(7.8)^0}{0!} - \frac{e^{-7.8}(7.8)^1}{1!} = 1 - .0004097 - .003200 = .9964$$

Using the R-function dpois, $P(X \geq 2) = 1 - dpois(0, 7.8) - dpois(1, 7.8) = 1 - .0004097 - .003200 = .9964$ or

Using the R-function ppois, $P(X \geq 2) = 1 - P(X \leq 1) = 1 - ppois(1, 7.8) = 1 - .0036 = .9964$

(5.) (10 points)

(a.) A has a t-distribution with df = 3 (A is the ratio of a N(0,1) r.v. and the square root of a Chi-square r.v. divided by its df. with the numerator and denominator r.v's having independent distributions)

(b.) B has a Cauchy distribution with location = 0 and scale =1 (B is the ratio of two independent N(0,1) r.v.'s)

(c.) C has a chi-squared distribution with df = 3 (C is the sum of independent squared N(0,1) r.v.s)

(d.) D has an F-distribution with $df_1 = 4,\ df_2 = 3$ (An F-distribution is the ratio of two independent Chi-square r.v.'s divided by their df's.)

(e.) E has an F-distribution with $df_1 = 1,\ df_2 = 3$ (An F-distribution is the ratio of two independent Chi-square r.v.'s divided by their df's.)

(6.) (10 points) Let U = .26 be a realization from a Uniform on (0,1) distribution.

(a.) W = Weibull(γ=4,α=1.5): $Q(u) = 1.5[-log(1-u)]^{1/4} \Rightarrow$

$W = Q(.26) = 1.5[-log(1-.26)]^{1/4} = 1.111$

(b.) $N = NegBin(r = 8, p = 0.7)$. Recall that the R functions for Negative Binomial are modeling the number of failures. Using the R function **pnbinom(x,8,.7)** with

x=c(0,1,2,3), we obtain the cdf, F(x), for X equal to the number failures before the 8th success:

$$F(x) = \begin{cases} 0.05764801 & x = 0 \\ 0.19600323 & x = 1 \\ 0.38278279 & x = 2 \\ 0.56956234 & x = 3 \end{cases}$$

Thus, with U=.26, we obtain $X = 2$ because $F(1) = .196 < .26 < .383 = F(2)$.

Therefore, N, the number of trials before the 8th success, $N = X + 8 = 2 + 8 = 10$.

(c.) B = Bin(20,.4): Using the R function **pbinom(x,20,.4)** with

x=c(5,6,7), we obtain

$$F(x) = \begin{cases} .1256 & x = 5 \\ .2500 & x = 6 \\ .4159 & x = 7 \end{cases}$$

Thus, with U=.26, we obtain $B = 7$ because $F(6) = .25 < .26 < .4159 = F(7)$

(d.) P = Poisson(λ=3): Using the R function **ppois(x,3)** with x=c(0,1,2), we obtain

$$F(x) = \begin{cases} .04978707 & x = 0 \\ .19914827 & x = 1 \\ .42319008 & x = 2 \end{cases}$$

Thus, with U=.26, we obtain $P = 2$ because $F(1) = .1991 < .26 < .4232 = F(2)$.

(e.) Y = Uniform on (0.3,2.5). Then the pdf is $f(y) = 1/(2.5-.3)$ for $.3 < y < 2.5$; 0 otherwise. Therefore, the cdf is given by

$$F(y) = 0 \text{ for } y \leq .3;\ F(y) = 1 \text{ for } y \geq 2.5;\ \text{For } .3 < y < 2.5,\ F(y) = \int_{.3}^{y} \frac{1}{2.5-.3}dy = \frac{1}{2.5-.3}(y-.3)$$

Let $u = F(y_u) = \frac{1}{2.5-.3}(y_u - .3)$, then solve for y_u yields $Q(u) = y_u = .3 + (2.5-.3)u$.

Therefore, with $U = .26$, $Y = .3 + (2.5-.3)(.26) = 0.872$

(7.) (30 points)

(a.) Hypergeometric - Sampling from finite population in which there are two types of units

(b.) Exponential - Distance between events in a Poisson process

(c.) Binomial (assuming 100,000 is very large) or Hypergeometric - Sampling from finite population in which there are two types of units

(d.) Poisson - counting the number of cracks larger than 20mm in a randomly selected pipe

(e.) Beta- Values of p are within (0,1) and have a skewed distribution

(f.) Weibull - Modeling extremes, maximum daily ozone level

(g.) Poisson - number of events occurring in a fixed length of pipe

(h.) Poisson - number of events occurring in a large number of genes

(i.) Binomial - number of occurrences of independent events in a fixed number of trials

(j.) Negative Binomial - number of trials until success, 100 subjects are accepted into the study

(k.) Cauchy - symmetric distribution with a large number of extreme values

(l.) Poisson - recording number of events in space - cracks on wing

(m.) Gamma - Time until 15th event in a Poisson process

(n.) Binomial - Counting number of trials (cows) in which a success occurs (5 or fewer ticks)

(o.) Hypergeometric - Sampling from finite population in which there are two types of units