Stat 641

Solutions for Assignment 2

(1.) (10 points) In the following expressions let y be a non-negative integer. Using the expression $\sum_{k=0}^{m} ab^k = a\frac{1-b^{m+1}}{1-b}$, we have $F(y) = P[Y \le y] = \sum_{k=0}^{y} p(1-p)^k \Rightarrow$

$$(a.)\ F(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1-(1-p)^{y+1} & \text{if } y \ge 0 \end{cases} \qquad (b.)\ Q(u) = \begin{cases} 0 & \text{if } u = 0 \\ inf\{y : y \ge \frac{log(1-u)}{log(1-p)} - 1\} & \text{if } 0 < u \le 1 \end{cases}$$

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

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

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

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.) $P(Y > 30) = 1 - P(Y \le 30) = 1 - F(30) = e^{-\left(\frac{30-10}{25}\right)^2} = .5273$

(d.) $Q(.4) = 10 + 25(-log(1-.4))^{1/2} = 27.868$

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

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

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

(b.) An examination of the expression for Weibull pdf with parameters $(\alpha,\ \gamma)$, suggests $\alpha = \beta^{1/\gamma}$ as the appropriate scale parameter. The pdf becomes

$$f(y) = \frac{\gamma}{\alpha^\gamma} y^{\gamma-1} e^{-y^\gamma/\alpha^\gamma} = \frac{\gamma}{\alpha}(y/\alpha)^{\gamma-1} e^{-(y/\alpha)^\gamma}$$

.

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 $\lambda = 0.25$. Then

$$P(X \ge 1) = 1 - P(X = 0) = 1 - \frac{e^{-0.25}(0.25)^0}{0!} = 0.221 = 1 - dpois(0, .25) \qquad (\underline{\text{dpois}} \text{ is R-function})$$

(b) Let Y be the number of emissions in a year. Y has a Poisson distribution with $\lambda = 0.25 \times 52 = 13$. Then

$$P(Y \ge 1) = 1 - P(Y = 0) = 1 - \frac{e^{-13}(13)^0}{0!} = .9999977 = 1 - dpois(0, 13)$$

(5.) (10 points) (a.) R has a chi-squared distribution with df = 4

(b.) W has a t-distribution with df = 6 (W 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)

(c.) Y has an F-distribution with $df_1 = 1,\ df_2 = 7$ (F is the ratio of two independent Chi-square r.v.'s divided by their df's.)

(d.) T has a Cauchy distribution with location = 0 and scale =1 (T is the ratio of two independent N(0,1) r.v.'s)

(e.) S has an F-distribution with $df_1 = 2,\ df_2 = 3$ (F is the ratio of two independent Chi-square r.v.'s divided by their df's.)

(6.) (10 points) Let U = .38 be a realization from a Uniform on (0,1) distribution.

(a.) W = Weibull(γ=4,α=1.5): $Q(u) = 1.5[-log(1-u)]^{1/4} \Rightarrow$
$W = Q(.38) = 1.5[-log(1-.38)]^{1/4} = 1.247$

(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=.38, we obtain $X = 2$ because $F(1) = .196 < .38 < .383 = F(2)$. 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=.38, we obtain $B = 7$ because $F(6) = .25 < .38 < .4159 = F(7)$

(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=.38, we obtain $P = 2$ because $F(1) = .1991 < .38 < .4232 = F(2)$.

(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

$F(y) = 0$ for $y \le .3$; $F(y) = 1$ for $y \ge 2.5$; 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 $U = .38$, $Y = .3 + (2.5 - .3)(.38) = 1.136$

(7.) (30 points)

(a.) F distribution - Ratio of independent chi-square r.v.s

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

(c.) Cauchy - Symmetric with heavy tails

(d.) Weibull - Hazard function is a power function

(e.) Hypergeometric - Sampling without replacement from a fixed population

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

(g.) Weibull - Hazard function is a power function

(h.) Binomial - Assuming that diode failures are independent, we have 200 independent Bernoulli trials with same probability (.001) of success (diode fails) on each trial

(i.) Hypergeometric - Sampling without replacement from a fixed population

(j.) Normal - $P[\mu-\sigma < X < \mu+\sigma] = .7 \approx .68, P[\mu-2\sigma < X < \mu+2\sigma] = .95 \approx .9545, P[|X-\mu| > 3\sigma] \approx 0$

(k.) Negative binomial - Observing independent Bernoulli trials until 50th success

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

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

(n.) Chi-square - Sum of 10 squared standard normal r.v.s

(o.) Beta - distribution on (0,1) which can be right or left skewed