UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Basic Exam - Probability
Wednesday, August 19, 2020

Show all work in your solution to each problem. 60 points are needed to pass at the Masters Level and 75 to pass at the Ph.D. level.

1. Suppose that $Z \sim$ Bernoulli(p), and that $P(X = 0 \mid Z = 0) = 1$ and that given $Z = 1$, $X \sim \text{Gamma}(\alpha, \beta)$. (This is known as a zero-inflated model.) Recall that the expectation of a Gamma random variable is α/β and the variance is α/β^2.

   (a) Find $E[X]$.

   (b) Find $\text{Var}(X)$.

2. Let $D \in \{0, 1\}$ denote unknown disease status of a person. Suppose that a diagnostic test Y for detecting D is Bernoulli distributed.

   (a) Suppose that $P(Y = 1 \mid D = 0) = 0.1$ and $P(Y = 1 \mid D = 1) = 0.8$ and that the overall prevalence of the disease in the population is $P(D = 1) = 0.01$. Find the probability that someone has the disease given that they test positive (known as the positive predictive value of the test) and the probability that someone does not have the disease given that they test negative (known as the negative predictive value of the test).

   Suppose now that we observe n IID observations $(Y_1, D_1), \ldots, (Y_n, D_n)$ and that we wish to estimate $\theta = P(Y = 1 \mid D = 1) = p/q$ for $p = P(Y = 1, D = 1)$ and $q = P(D = 1)$.

   (b) Suppose that $q > 0$ is known and consider $\hat{\theta}_n = p_n/q$, where $p_n = \frac{1}{n}\sum_{i=1}^n Y_i D_i$. (i) Find $\text{Var}(\hat{\theta}_n)$ in terms of p, q, and n. (ii) Show that $\hat{\theta}_n$ converges in probability to θ.

   (c) Suppose now that $q > 0$ is unknown and consider $\theta_n^* = p_n/q_n$, where $q_n = \frac{1}{n}\sum_{i=1}^n D_i$. Show that θ_n^* converges in probability to θ.

   (d) Letting $p_n = \frac{1}{n}\sum_{i=1}^n Y_i D_i$ and $q_n = \frac{1}{n}\sum_{i=1}^n D_i$, show that

   $$\sqrt{n}\left[\begin{pmatrix} p_n \\ q_n \end{pmatrix} - \begin{pmatrix} p \\ q \end{pmatrix}\right] \to_d N_2\left(0, \Sigma\right),$$

   and find Σ in terms of p and q.

   (e) Using part (d), show that $\sqrt{n}(\theta_n^* - \theta)$ converges in distribution to $N(0, \sigma^2)$, and find σ^2 in terms of p and q. (Hint: use the delta method.)

3. Let X be a Uniform(0,1) random variable.

(a) Find a transformation $g()$ so that $Y = g(X)$ has an exponential distribution with rate $\lambda > 0$, i.e. has pdf $g(x) = \lambda \exp(-\lambda x)$. (Please also show that $g()$ is the right transformation.)

(b) Derive the moment generating function of Y.

(c) Prove that $Pr(Y > a + b|Y > b) = Pr(Y > a), a, b > 0$.

(d) Consider Y_1 and Y_2, *iid* exponentials with rate λ. Derive the distribution of $Z = \frac{Y_1}{Y_1+Y_2}$.

(e) Suppose that Y measures the time until a rare event, and the expected number of events in one year is λ. Show that the probability of no events in a year is approximately $1 - \lambda$, and the approximation is better when λ is close to zero.

4. A real-valued random variable X is said to follow the Pareto distribution with scale $\lambda \in (0, \infty)$ and shape $\alpha \in (0, \infty)$ if it has distribution function $F(x; \lambda, \alpha)$ given by

$$P(X \leq x; \lambda, \alpha) = F(x; \lambda, \alpha) = 1 - (\lambda/x)^{\alpha}$$

for $x > \lambda$, and 0 otherwise, and density function $f(x; \lambda, \alpha)$ given by

$$f(x; \lambda, \alpha) = \frac{\alpha\lambda^{\alpha}}{x^{\alpha+1}}$$

again **for** $x > \lambda$, and 0 otherwise.

Suppose that $X_1, X_2, \ldots, X_n$ is a sequence of IID Pareto(λ, α) random variables.

(a) Find the density function of $U_i = \log(X_i/\lambda)$, and show that $E[\log(X_i/\lambda)] = \alpha^{-1}$.

(b) Suppose that λ is known. Define the statistic $\hat{\alpha}_n = \left[\frac{1}{n}\sum_{i=1}^{n} \log(X_i/\lambda)\right]^{-1}$. Show that $\hat{\alpha}_n$ converges in probability to α.

(c) Show that $\sqrt{n}(\hat{\alpha}_n - \alpha)$ converges in distribution to a $N(0, \sigma^2)$ random variable, and find σ.

(d) (i) Find the distribution of $\hat{\lambda}_n = \min\{X_1, \ldots, X_n\}$, and (ii) show that $\hat{\lambda}_n$ converges in probability to λ (hint: show that $P(\hat{\lambda}_n \leq \lambda = 0)$ and $P(\hat{\lambda}_n > \lambda + \varepsilon) \to 0$ for all $\varepsilon > 0$).

5. Suppose X is normal with mean μ and variance σ^2. The pdf is $\phi(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}$, and you may use the notation $\Phi(x)$ for the cdf evaluated at x. Let $Y = [X|a < X < b]$.

(a) Find and draw the pdf of Y.

(b) Find the mean and variance of Y.