Part 1: Sample Problems for the Elementary Section of Qualifying Exam in Probability and Statistics

https://www.soa.org/Files/Edu/edu-exam-p-sample-quest.pdf

Part 2: Sample Problems for the Advanced Section of Qualifying Exam in Probability and Statistics

Probability

1. The *Pareto* distribution, with parameters α and β, has pdf

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}}, \ \alpha < x < \infty, \ \alpha > 0, \ \beta > 0.$$

(a) Verify that $f(x)$ is a pdf.
(b) Derive the mean and variance of this distribution.
(c) Prove that the variance does not exist if $\beta \leq 2$.

2. Let $U_i, \ i = 1, 2, \ldots,$ be independent uniform(0,1) random variables, and let X have distribution

$$P(X = x) = \frac{c}{x!}, \ x = 1, 2, 3, \ldots,$$

where $c = 1/(e-1)$. Find the distribution of $Z = min\{U_1, \ldots, U_X\}$.

3. A point is generated at random in the plane according to the following polar scheme. A radius R is chosen, where the distribution of R^2 is χ^2 with 2 degrees of freedom. Independently, an angle θ is chosen, where $\theta \sim$ uniform$(0, 2\pi)$. Find the joint distribution of $X = R\cos\theta$ and $Y = R\sin\theta$.

4. Let X and Y be iid $N(0,1)$ random variables, and define $Z = \min(X, Y)$. Prove that $Z^2 \sim \chi_1^2$.

5. Suppose that $\mathcal{B}$ is a σ-field of subsets of Ω and suppose that $P: \ \mathcal{B} \to [0,1]$ is a set function satisfying:
(a) P is finitely additive on $\mathcal{B}$;
(b) $0 \leq P(A) \leq 1$ for all $A \in \mathcal{B}$ and $P(\Omega) = 1$;
(c) If $A_i \in \mathcal{B}$ are disjoint and $\cup_{i=1}^{\infty} A_i = \Omega$, then $\sum_{i=1}^{\infty} P(A_i) = 1$.
Show that P is a probability measure on $\mathcal{B}$ in Ω.

6. Suppose that $\{X_n\}_{n=1}^{\infty}$ is a sequence of i.i.d. random variables and c_n is an increasing sequence of positive real numbers such that for all $\alpha > 1$, we have

$$\sum_{n=1}^{\infty} P[X_n > \alpha^{-1} c_n] = \infty$$

and

$$\sum_{n=1}^{\infty} P[X_n > \alpha c_n] < \infty.$$

Prove that

$$P\left[\limsup_{n\to\infty} \frac{X_n}{c_n} = 1\right] = 1.$$

7. Suppose for $n \geq 1$ that $X_n \in L_1$ are random variables such that $sup_{n\geq 1} E(X_n) < \infty$. Show that if $X_n \uparrow X$, then $X \in L_1$ and $E(X_n) \to E(X)$.

8. Let X be a random variable with distribution function $F(x)$.
(a) Show that

$$\int_{\mathbb{R}} (F(x+a) - F(x))dx = a.$$

(b) If F is continuous, then $E[F(X)] = \frac{1}{2}$.

9. (a) Suppose that $X_n \xrightarrow{P} X$ and g is a continuous function. Prove that $g(X_n) \xrightarrow{P} g(X)$.
(b) If $X_n \xrightarrow{P} 0$, then for any $r > 0$,

$$\frac{|X_n|^r}{1 + |X_n|^r} \xrightarrow{P} 0$$

and

$$E[\frac{|X_n|^r}{1 + |X_n|^r}] \to 0.$$

10. Suppose that $\{X_n, n \geq 1\}$ are independent non-negative random variables satisfying $E(X_n) = \mu_n$, $\mathrm{Var}(X_n) = \sigma_n^2$. Define for $n \geq 1$, $S_n = \sum_{i=1}^{n} X_i$ and suppose that $\sum_{n=1}^{\infty} \mu_n = \infty$ and $\sigma_n^2 \leq c\mu_n$ for some $c > 0$ and all n. Show

$$\frac{S_n}{E(S_n)} \xrightarrow{P} 1.$$

11. (a) If $X_n \to X$ and $Y_n \to Y$ in probability, then $X_n + Y_n \to X + Y$ in probability.
(b) Let $\{X_i\}$ be iid, $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. Set $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$. Show that

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \to \sigma^2$$

in probability.

12. Suppose that the sequence $\{X_n\}$ is *fundamental in probability* in the sense that for ε positive there exists an N_ε such that $P[|X_n - X_m| > \varepsilon] < \varepsilon$ for $m,\ n > N_\varepsilon$.
(a) Prove that there is a subsequence $\{X_{n_k}\}$ and a random variable X such that $\lim_k X_{n_k} = X$ with probability 1 (i.e. almost surely).
(b) Show that $f(X_n) \to f(X)$ in probability if f is a continuous function.

Statistics

1. Suppose that $X = (X_1, \cdots, X_n)$ is a sample from the probability distribution P_θ with density

$$f(x|\theta) = \begin{cases} \theta(1+x)^{-(1+\theta)}, & \text{if} \quad x > 0 \\ 0, & \text{otherwise} \end{cases}$$

for some $\theta > 0$.
(a) Is $\{f(x|\theta), \theta > 0\}$ a one-parameter exponential family? (explain your answer).
(b) Find a sufficient statistic $T(X)$ for $\theta > 0$.

2. Suppose that $X_1, \cdots, X_n$ is a sample from a population with density

$$p(x, \theta) = \theta a x^{a-1} \exp(-\theta x^a),\ x > 0,\ \theta > 0,\ a > 0.$$

(a) Find a sufficient statistic for θ with a fixed.
(b) Is the sufficient statistic in part (a) minimally sufficient? Give reasons for your answer.

3. Let $X_1, \cdots, X_n$ be a random sample from a gamma(α, β) population.
(a) Find a two-dimensional sufficient statistic for (α, β).
(b) Is the sufficient statistic in part (a) minimally sufficient? Explain your answer.
(c) Find the moment estimator of (α, β).
(d) Let α be known. Find the best unbiased estimator of β.

4. Let $X_1, \ldots, X_n$ be iid Bernoulli random variables with parameter θ (probability of a success for each Bernoulli trial), $0 < \theta < 1$. Show that $T(X) = \sum_{i=1}^n X_i$ is minimally sufficient.

5. Suppose that the random variables $Y_1, \cdots, Y_n$ satisfy

$$Y_i = \beta x_i + \varepsilon_i, i = 1, \cdots, n,$$

where $x_1, \cdots, x_n$ are fixed constants, and $\varepsilon_1, \cdots, \varepsilon_n$ are iid $N(0, \sigma^2)$, σ^2 unknown.
(a) Find a two-dimensional sufficient statistics for (β, σ^2).
(b) Find the MLE of β, and show that it is an unbiased estimator of β.
(c) Show that $[\sum(Y_i/x_i)]/n$ is also an unbiased estimator of β.

6. Let $X_1, \cdots, X_n$ be iid $N(\theta, \theta^2)$, $\theta > 0$. For this model both $\bar{X}$ and cS are unbiased estimators of θ, where

$$c = \frac{\sqrt{n-1}\Gamma((n-1)/2)}{\sqrt{2}\Gamma(n/2)}.$$

(a) Prove that for any number a the estimator $a\bar{X} + (1-a)(cS)$ is an unbiased estimator of θ.
(b) Find the value of a that produces the estimator with minimum variance.
(c) Show that $(\bar{X}, S^2)$ is a sufficient statistic for θ but it is not a complete sufficient statistic.

7. Let $X_1, \cdots, X_n$ be i.i.d. with pdf

$$f(x|\theta) = \frac{2x}{\theta} \exp\{-\frac{x^2}{\theta}\}, \ x > 0, \ \theta > 0.$$

(a) Find the Fisher information

$$I(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\mathbf{X}|\theta)\right)^2\right],$$

where $f(\mathbf{X}|\theta)$ is the joint pdf of $\mathbf{X} = (X_1, \ldots, X_n)$.
(b) Show that $\frac{1}{n}\sum_{i=1}^n X_i^2$ is an UMVUE of θ.

8. Let $X_1, \cdots, X_n$ be a random sample from a $n(\theta, \sigma^2)$ population, σ^2 known. Consider estimating θ using squared error loss. Let $\pi(\theta)$ be a $n(\mu, \tau^2)$ prior distribution on θ and let δ^π be the Bayes estimator of θ. Verify the following formulas for the risk

function, Bayes estimator and Bayes risk.
(a) For any consatnts a and b, the estimator $\delta(X) = a\bar{X} + b$ has risk function

$$R(\theta, \delta) = a^2 \frac{\sigma^2}{n} + (b - (1-a)\theta)^2.$$

(b) Show that the Bayes estimator of θ is given by

$$\delta^\pi(X) = E(\theta|\bar{X}) = \frac{\tau^2}{\tau^2 + \sigma^2/n}\bar{X} + \frac{\sigma^2/n}{\tau^2 + \sigma^2/n}\mu.$$

(c)Let $\eta = \sigma^2/(n\tau^2 + \sigma^2)$. The risk function for the Bayes estimator is

$$R(\theta, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2(\theta - \mu)^2.$$

(d) The Bayes risk for the Bayes estimator is

$$B(\pi, \delta^\pi) = \tau^2 \eta.$$

9. Suppose that $X = (X_1, \cdots, X_n)$ is a sample from normal distribution $N(\mu, \sigma^2)$ with $\mu = \mu_0$ known.
(a) Show that $\hat{\sigma_0}^2 = n^{-1}\sum_{i=1}^n (X_i - \mu_0)^2$ is a uniformly minimum variance unbiased estimate (UMVUE) of σ^2.
(b) Show that $\hat{\sigma_0}^2$ converges to σ^2 in probability as $n \to \infty$.
(c) If μ_0 is not known and the true distribution of X_i is $N(\mu, \sigma^2)$, $\mu \neq \mu_0$, find the bias of $\hat{\sigma_0}^2$.

10. Let $X_1, \cdots, X_n$ be i.i.d. as $X = (Z, Y)^T$, where $Y = Z + \sqrt{\lambda} W$, $\lambda > 0$, Z and W are independent $N(0, 1)$.
(a)Find the conditional density of Y given $Z = z$.
(b)Find the best predictor of Y given Z and calculate its mean squared prediction error (MSPE).
(c)Find the maximum likelihood estimate (MLE) of λ.
(d)Find the mean and variance of the MLE.

11. Let $X_1, \cdots, X_n$ be a sample from distribution with density

$$p(x, \theta) = \theta x^{\theta - 1} 1\{x \in (0, 1)\}, \ \theta > 0.$$

(a) Find the most powerful (MP) test for testing $H : \theta = 1$ versus $K : \theta = 2$ with $\alpha = 0.10$ when $n = 1$.

(b) Find the MP test for testing $H : \theta = 1$ versus $K : \theta = 2$ with $\alpha = 0.05$ when $n \geq 2$.

12. Let $X_1, \cdots, X_n$ be a random sample from a $N(\mu_1, \sigma_1^2)$, and let $Y_1, \cdots, Y_m$ be an independent random sample from a $N(\mu_2, \sigma_2^2)$. We would like to test

$$H : \mu_1 = \mu_2 \quad \text{versus} \quad K : \mu_1 \neq \mu_2$$

with the assumption that $\sigma_1^2 = \sigma_2^2$.
(a) Derive the likelihood ratio test (LRT) for these hypotheses. Show that the LRT can be based on the statistic

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2 \left(\frac{1}{n} + \frac{1}{m}\right)}},$$

where

$$S_p^2 = \frac{1}{n+m-2}\left(\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2\right).$$

(b)Show that, under H, T has a t_{n+m-2} distribution.

Part 3: Required Proofs for Probability and Statistics Qualifying Exam

In what follows X_i's are always i.i.d. real random variables (unless otherwise specified).

You are allowed to use some well known theorems (like Lebesgue Dominant Convergence Theorem or Chebyshev inequality), but you must state them and explain how and where do you use them.

Warning: If X and Y have the same moment generating function it does not mean that their distributions are the same.

1. Prove that

$$\text{if } X_n \to X_0 \text{ in probability, then } X_n \to X_0 \text{ in distribution.}$$

Offer a connterexample for the converse.

2. Prove that

$$\text{if } E|X_n - X_0| \to 0. \quad \text{then } X_n \to X_0 \text{ in probability.}$$

Offer a connterexample for the converse.

3. We define $d_{BL}(X_n, X_0) = Sup_{H\in BL}|EH(X_n) - EH(X_0)|$, where BL is a set of all real functions that are Lipshitz and bounded by 1. Prove that

$$\text{if } d_{BL}(X_n, X_0) \to 0, \quad \text{then } P(X_n \le t) \to P(X_0 \le t)$$

for every t for which function $F(t) = P(X_0 \le t)$ is continuous.

4. Prove that

$$\text{if } X_n \to X_0 \text{ in probability and } Y_n \to Y_0 \text{ in distribution,}$$

then

$$X_n + Y_n \to X_0 + Y_0 \text{ in distribution.}$$

5. Prove that if $EX_i^2 < \infty$, then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to E(X_1) \text{ in probability.}$$

6. (Count as two) Prove that if $E(|X_i|)$ exists, then

$$n^{-1}\sum_{i=1}^{n} X_i \to EX_1 \text{ in probability.}$$

7. Prove that if $EX_i^4 < \infty$, then

$$n^{-1}\sum_{i=1}^{n} X_i \to EX_1 \text{ a.s.}$$

Hint: Work with: $P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} |n^{-1}\sum_{i=1}^{n} X_i - EX_1| > \varepsilon)$.

8. (Count as two) Prove that if $E|X_i|^3 < \infty$, then

$$n^{-1/2}\sum_{i=1}^{n}(X_i - EX_1) \to Z \text{ in distribution,}$$

where Z is a centered normal random variable with $E(Z^2) = Var(X_i) = \sigma^2$.

9. Prove: For any $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$

$$E|XY| \le (E|X|^p)^{1/p}\,(E|X|^q)^{1/q}\,.$$

10. Prove that if

$$X_n \to X_0 \text{ in probability and } |X_i| \le M < \infty,$$

then

$$E|X_n - X_0| \to 0.$$

11. (Count as two) Let $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le t\}}$ and $F(t) = P(X_i \le t)$ be a continuous function. Then

$$\sup_t |F_n(t) - F(t)| \to 0 \text{ in probability.}$$

12. Let X and Y be independent Poisson random variables with their parameters equal λ. Prove that $Z = X + Y$ is also Poisson and find its parameter.

13. Let X and Y be independent normal random variables with $E(X) = \mu_1, E(Y) = \mu_2, Var(X) = \sigma_1^2, Var(Y) = \sigma_2^2$. Show that $Z = X + Y$ is also normal and find $E(Z)$ and $Var(Z)$.

14. Let X_n converge in distribution to X_0 and let $f : R \to R$ be a continuous function. Show that $f(X_n)$ converges in distribution to $f(X_0)$.

15. Using only the Axioms of probability and set theory, prove that
a)
$$A \subset B \Rightarrow P(A) \leq P(B).$$

b)
$$P(X + Y > \varepsilon) \leq P(X > \varepsilon/2) + P(Y > \varepsilon/2).$$

c) If A and B are independent events, then A^c and B^c are independent as well.
d) If A and B are mutually exclusive and $P(A) + P(B) > 0$, show that
$$P(A|A \cup B) = \frac{P(A)}{P(A) + P(B)}.$$

16. Let A_i be a sequence of events. Show that
$$P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

17. Let A_i be a sequence of events such that $A_i \subset A_{i+1}$, $i = 1, 2, ...$ Prove that
$$\lim_{n\to\infty} P(A_n) = P(\cup_{i=1}^{\infty} A_i).$$

18. Formal definition of weak convergence states that $X_n \to X_0$weakly if for every continuous and bounded function f :$R \to R$, $Ef(X_n) \to Ef(X_0)$. Show that:
$$X_n \to X_0 \text{ weakly } \Rightarrow P(X_n \leq t) \to P(X \leq t)$$
for every t for which the function $F(t) = P(X \leq t)$ is continuous.

19. (Borel-Cantelli lemma). Let A_i be a sequence of events such that $\sum_{i=1}^{\infty} P(A_i) < \infty$, then
$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = 0.$$

20. Consider the linear regression model $Y = X\beta + e$, where Y is an $n \times 1$ vector of the observations, X is the $n \times p$ design matrix of the levels of the regression variables, β is an $p \times 1$ vector of the regression coefficients, and e is an $n \times 1$ vector of random errors. Prove that the least squares estimator for β is $\hat{\beta} = (X'X)^{-1}X'Y$.

21. Prove that if X follows a F distribution $F(n_1, n_2)$, then X^{-1} follows $F(n_2, n_1)$.

22. Let $X_1, \cdots, X_n$ be a random sample of size n from a normal distribution $N(\mu, \sigma^2)$. We would like to test the hypothesis $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. When σ is known, show that the power function of the test with type I error α under true population mean $\mu = \mu_1$ is $\Phi(-z_{\alpha/2} + \frac{|\mu_1 - \mu_0|\sqrt{n}}{\sigma})$, where $\Phi(.)$ is the cumulative distribution function of a standard normal distribution and $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

23. Let $X_1, \cdots, X_n$ be a random sample of size n from a normal distribution $N(\mu, \sigma^2)$. Prove that (a) the sample mean $\bar{X}$ and the sample variance S^2 are independent; (b) $\frac{(n-1)S^2}{\sigma^2}$ follows a Chi-squared distribution $\chi^2(n-1)$.

Qualifying Exam on Probability and Statistics
August 24, 2017

Name:

Instruction: There are ten problems at two levels: 5 problems at elementary level and 5 proof problems at graduate level. Therefore, please make sure to budget your time to complete problems at both levels. Show your detailed steps in order to receive credits. You have 3 hours to complete the exam. GOOD LUCK!

Level 1: Elementary Problems

1. A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not actually present. One percent of the population actually has the disease.

   (a). What is the probability that the test indicates the presence of the disease?

   (b). Calculate the probability that a person actually has the disease given that the test indicates the presence of the disease.

2. Let X and Y be independent and identically distributed random variables such that the moment generating function of $X+Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t}, -\infty < t < \infty.$$

(a). Find $P(X \leq 0)$.

(b). Find $\mu = E(X)$ and $\sigma^2 = Var(X)$.

3. On May 5, in a certain city, temperatures have been found to be normally distributed with mean $\mu = 24$°C and variance $\sigma^2 = 9$. The record temperature on that day is 27°C.

   (a). What is the probability that the record of 27°C will be broken on next May 5 ?

   (b). What is the probability that the record of 27°C will be broken at least 3 times during the next 5 years on May 5 ? (Assume that the temperatures during the next 5 years on May 5 are independent.)

   (c). How high must the temperature be to place it among the top 5% of all temperatures recorded on May 5 ?

4. A company offers earthquake insurance. Annual premiums are modeled by an exponential random variable with mean 2. Annual claims are modeled by an exponential random variable with mean 1. Premiums and claims are independent. Let X denote the ratio of claims to premiums. Determine the probability density function $f(x)$ of X.

5. Ten cards from a deck of playing cards are in a box: two diamonds, three spades, and five hearts. Two cards are randomly selected without replacement. Find the conditional variance of the number of diamonds selected, given that no spade is selected.

Level 2: Proof Problems

1. Let $\{A_i\}$ be a sequence of events satisfying $\sum_{i=1}^{\infty} P(A_i) < \infty$. Show that

$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = 0.$$

2. Let $\{X_i\}_{i=1}^n$ be i.i.d. and normally distributed with mean μ and variance σ^2. Show that

   (a). the sample mean $\bar{X}_n$ and sample variance S_n^2 are independent.

   (b). $\frac{(n-1)S_n^2}{\sigma^2}$ follows a chi-squared distribution χ^2_{n-1}.

3. Show that the sequence of random variables X_1, $X_2, \dots$ converges in probability to a constant a if and only if the sequence also converges in distribution to a. That is the statement

$$P(|X_n - a| > \varepsilon) \to 0 \text{ for every } \varepsilon > 0$$

is equivalent to

$$P(X_n \leq x) \to \begin{cases} 0 & \text{if } x < a \\ 1 & \text{if } x > a. \end{cases}$$

Questions 4-5: Let $X_1, \dots, X_n$ be a simple random sample of size n from a population X, and $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ be the sample mean and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ be the sample variance, respectively.

4. If $E(X) = \mu$ and $0 < Var(X) = \sigma^2 < \infty$, show that

   (a). $S_n^2 \to \sigma^2$ in probability as $n \to \infty$.

   (b). $\frac{S_n}{\sigma} \to 1$ in probability as $n \to \infty$.

5. Assume that X is normally distributed with both mean and standard deviation being θ, $\theta > 0$.

   (a). Show that both $\bar{X}_n$ and cS_n are unbiased estimators of θ, where

   $$c = \frac{\sqrt{n-1}\Gamma((n-1)/2)}{\sqrt{2}\Gamma(n/2)}.$$

   (b). Find the value of t so that the linear interpolation estimator $t\bar{X}_n + (1-t)(cS_n)$ has minimum variance.

   (c). Show that $(\bar{X}_n, S_n^2)$ is a sufficient statistic for θ but it is not a complete sufficient statistic.

Qualifying Exam on Probability and Statistics
January 13, 2017

CHALLENGING PART

1) Let X_n and S_n be two sequences of real random variables such that S_n converges to Z in distribution, where Z is a continuous random variable. If X_n converges to a constant K in probability, show that the following is true

$$X_n S_n \to KZ \text{ in distribution as } n \to \infty$$

Hint: First show that one can assume (without loss of generality) that $|X_n| < M$ for sufficiently large M.

2) Let X_n be a sequence of random variables such that $EX_n = \mu$ and $|Cov(X_n, X_m)| \leq \frac{1}{1+|n-m|}$ for all $n, m \in N$.

a) Show that

$$n^{-1} \sum_{i=1}^{n} X_i \to \mu \text{ in probability as } n \to \infty$$

Hint: Chebishev.

b) Show that

$$n^{-1/2} \sum_{i=1}^{n} (X_i - \mu) \text{ does not converge in distribution (as } n \to \infty)$$

Hint: $Var(n^{-1/2} \sum_{i=1}^{n} X_i) \to \infty$.

3) Let (X, Y) be a random vector and let $H(s,t) = P(X \leq s, Y \leq t)$, $F(s) = P(X \leq s)$, $G(t) = P(Y \leq t)$. We also assume that F and G are continuous functions and H is not necessarily continuous. Define $C(a,b) = H(F^{-1}(a), G^{-1}(b))$, for $(a,b) \in [0,1]^2$.

a) Show that
$$H(s,t) = C(F(s), G(t))$$

b) Show that for every $s \leq s_o$ and $t \leq t_o$ the following is true

$$C(s,t) \leq C(s_o, t_o)$$

PROOF PART

Prove the following:

1) Using only the Axioms of probability and set theory, prove that

a)

$$A \subset B \Rightarrow P(A) \leq P(B)$$

b)

$$P(X + Y > \varepsilon) \leq P(X > \varepsilon/2) + P(Y > \varepsilon/2)$$

c) If A and B are independent events than A^c and B^c are independent as well

d) If A and B are mutually exclusive and $P(A) + P(B) > 0$ than show that

$$P(A/A \cup B) = \frac{P(A)}{P(A) + P(B)}$$

2) If

$$X_n \to X_0 \text{ in probability and } |X_i| \leq M < \infty$$

then

$$E|X_n - X_0| \to 0$$

3) This problem count as two. If EX_i exists then

$$n^{-1} \sum_{i=1}^{n} X_i \to EX_1 \text{ in probability}$$

Qualifying Exam on Probability and Statistics
Spring, January 19, 2016

Instruction: You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are three parts in the exam. Please budget your time wisely for all three parts. There are 10 problems in Elementary part, 3 problems in Challenging part and 3 proof problems. The suggested passing grade for the three parts are: Elementary part 80%, Challenging part 50% and Proofs 80%.

1 Elementary part

(1). The number of injury claims per month is modeled by a random variable N with $P(N = n) = \frac{1}{(n+1)(n+2)}$ for non negative integral $n's$. Calculate the probability of at least one claim during a particular month, given that there have been at most four claims during that month.

(2). Let X be a continuous random variable with density function

$$f(x) = \frac{|x|}{10} \text{ for } x \in [-1, 4] \text{ and } \ f(x) = 0 \text{ otherwise.}$$

Calculate $E(X)$.

(3). A device that continuously measures and records sesmic activity is placed in a remote region. The time to failure of this device, T, is exponentialy distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T, 2)$. Calculate $E(X)$.

(4). The time until failure, T, of a product is modeled by uniform distribution on $[0, 10]$. An extended warranty pays a benefit of 100 if failure occurs between $t = 1.5$ and $t = 8$. The present value, W of this benefit is

$$W = 100e^{-0.04T} \text{ for } T \in [1.5, 8] \text{ and zero otherwise.}$$

Calculate $P(W < 79)$.

(5). On any given day, a certain machine has either no malfunctions or exactly one malfunction. The probability of malfunction on any given day is 0.4. Machine malfunctions on different days are mutually independent. Calculate the probability that the machine has its third malfunction on the fifth day, given that the machine has not had three malfunctions in the first three days.

(6). Two fair dice are rolled. Let X be the absolute value of the difference between the two numbers on the dice. Calculate $P(X < 3)$.

(7). A driver and a passenger are in a car accident. Each of them independently has probability 0.3 of being hospitalized. When a hospilatization occurs, the loss is uniformly distributed on $[0, 1]$. When two hospitalization occur, the losses are independent. Calculate the expected number of people in the car who are hospitalized, given that the total loss due to hospitalization is less than 1.

(8). Let X and Y be independent and identically distributed random variables such that the moment generating function for $X + Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t} \text{ for } \ t \in (-\infty, \infty)$$

Calculate $P(X \leq 0)$.

(9). The number of workplace injuries, N, occuring in a factory on any given day is Poisson distributed with mean λ. The parameter λ itself is a random variable that is determined by the level of activity in the factory and is uniformly distributed on inteval $[0, 3]$. Calculate $Var(N)$.

(10). Let X and Y be continuous random variables with joint density function

$$f(x, y) = \begin{cases} 24xy, & \text{for } \ 0 < y < 1 - x, \ x \in (0, 1); \\ 0, & \text{otherwise} \end{cases}$$

Calculate $P(Y < X | X = \frac{1}{3})$.

2 Challenging Part

(1). Let Y be a non negative random variable. Show that

$$EY \leq \sum_{k=0}^{\infty} P(Y > k) \leq EY + 1.$$

(2). Let X_n be a sequence of random variables such that $\sqrt{n}(X_n - \mu) \to N(0, \sigma^2)$ in distribution. For any given function g and a specific μ, suppose that $g'(\mu)$ exists and $g'(\mu) \neq 0$. Then prove that

$$\sqrt{n}(g(X_n) - g(\mu)) \to N(0, \sigma^2[g'(\mu)]^2) \ \text{ in distribution.}$$

(3). Let $\{X_n\}$ be a sequence of random variables with $E(X_n) = 0$, and $Var(X_n) \leq C$ (C is a constant), $E(X_i X_j) \leq \rho\big(i - j\big)$ for any $i > j$ and $\rho(n) \to 0$ as $n \to \infty$. Show that

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0 \ \text{ in probability.}$$

3 Proofs

(1). Let $\{X_n\}$ be a sequence of independent and identically distributed random variables with $E|X_n| < \infty$. Prove or disprove the following statement

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to EX_1 \quad \text{in probability as } \ n \to \infty.$$

(2). Let $X_n : \Omega \to R^d$ and such that X_n converges weakly (in distribution) to random vector Z. Let $F : R^d \to R$ be a continuous function and let $Y_n = F(X_n)$. Then prove or disprove the following statement:

$$Y_n \to F(Z) \text{ weakly } \ \text{(in distribution) as } \quad n \to \infty.$$

(3). Consider the linear regression model $Y = X\beta + e$, where Y is an $n \times 1$ vector of the observations, X is the $n \times p$ design matrix of the levels of the regression variables, β is a $p \times 1$ vector of regression coefficients and e is an $n \times 1$ vector of random errors. Show that the least square estimator for β is $\widehat{\beta} = (X'X)^{-1}X'Y$.