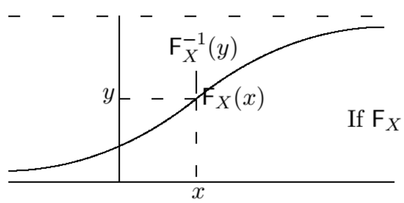
EE5137 Stochastic Processes: Problem Set 3
Assigned: 28/01/22, Due: Never

All problems here are optional. Please use these practice problems to prepare for the quiz.

1. Exercise 1.28 (Gallager's book) Suppose the rv X is continuous and has the CDF $F_X(x)$. Consider another rv $Y = F_X(X)$. That is, for each sample ω such that $X(\omega) = x$, we have $Y(\omega) = F_X(x)$. Show that Y is uniformly distributed in the interval 0 to 1.
   Solution For simplicity, first assume that $F_X(x)$ is strictly increasing in x, thus having the following appearance:

   Since $F_X(x)$ is continuous in x and strictly increasing from 0 to 1, there must be an inverse function $F_X^{-1}(x)$ such that for each $y \in (0,1), F_X^{-1}(y) = x$ for that x such that $F_X(x) = y$. For this y, then, the event $\{F_X(X) \le y\}$ is the same as the event $\{X \le F_X^{-1}(y)\}$. This is illustrated in the above figure. Using this equality for the given y,

   $$\begin{aligned} \Pr\{Y \le y\} &= \Pr\{F_X(X) \le y\} = \Pr\{X \le F_X^{-1}(y)\} && (1) \\ &= F_X(F_X^{-1}(y)) = y, && (2) \end{aligned}$$

   where in the final equation, we have used the fact that F_X^{-1} is the inverse function of F_X. This relation, for all $y \in (0,1)$, shows that Y is uniformly distributed between 0 and 1.

   If F_X is not strictly increasing, i.e., if there is any interval over which $F_X(x)$ has a constant value y, then we can define $F_X^{-1}(y)$ to have any given value within that interval. The above argument then still holds, although F_X^{-1} is no longer the inverse of F_X.

   If there is any discrete point, say z at which $\Pr\{X = z\} > 0$, then $F_X(x)$ cannot take on values in the open interval between $F_X(z) - a$ and $F_X(z)$ where $a = \Pr\{X = z\}$. Thus F_X is uniformly distributed only for continuous rv's.

2. Exercise 1.32 (Gallager's book) (The one-sided Chebyshev inequality) This inequality states that if a zero-mean rv X has a variance σ^2, then it satisfies the inequality

   $$\Pr\{X \ge b\} \le \frac{\sigma^2}{\sigma^2 + b^2} \quad \text{for every} \quad b > 0, \tag{3}$$

   which equality for some b only if X is binary and $\Pr\{X = b\} = \sigma^2/(\sigma^2 + b^2)$. We prove this here using the same approach as in Excercise 1.31. Let X be a non-zero rv that satisfies $\Pr\{X \ge b\} = \beta$ for some given $b > 0$ and $0 < \beta < 1$. The variance σ^2 of X can be expressed as

$$\sigma^2 = \int_{-\infty}^{b-} x^2 f_X(x)dx + \int_{b}^{\infty} x^2 f_X(x)dx. \tag{4}$$

We will first minimize σ^2 over all zero-mean X satisfying $\Pr\{X \geq b\} = \beta$.

(a) Show that the second integral in (4) satisfies $\int_b^\infty x^2 f_X(x)dx \geq b^2\beta$.

(b) Show that the first integral in (4) is constrained by

$$\int_{-\infty}^{b-} f_X(x)dx = 1-\beta, \quad \text{and} \quad \int_{-\infty}^{b-} x f_X(x)dx \leq -b\beta. \tag{5}$$

(c) Minimize the first integral in (4) subject to the constraint in (b). Hint: If you scale $f_X(x)$ up by $1/(1-\beta)$, it integrates to 1 over $(-\infty, b)$ and the second constraint becomes an expectation. You can then minimize the first integral in (4) by inspection.

(d) Combine the results in a) and c) to show that $\sigma^2 \geq b^2\beta/(1-\beta)$. Find the minimizing distribution. Hint: It is binary.

(e) Use (d) to establish (3). Also show (trivially) that if Y has a mean $\bar{Y}$ and variance σ^2, then $\Pr\{Y - \bar{Y} \geq b\} \leq \sigma^2/(\sigma^2 + b^2)$.

Solution:

(a) We first explain why we started trying to establish an upper bound to $\Pr\{X \geq b\}$ and then switched to minimizing the variance such that $\bar{X} = 0$ and $\Pr\{X \geq b\} = \beta$ for some given β. We will find that $\sigma^2_{\min} = b^2\beta/(1-\beta)$, or, equivalently, $\beta = \sigma^2_{\min}/(b^2 + \sigma^2_{\min})$. Thus for each zero-mean X satisfying $\Pr\{X \geq b\} = \beta$, we have (4), i.e.,

$$\Pr\{X \geq b\} = \beta = \frac{\sigma^2_{\min}}{\sigma^2_{\min} + b^2} \leq \frac{\sigma^2}{\sigma^2 + b^2}. \tag{6}$$

Solving part (a),

$$\int_b^\infty x^2 f_X(x) \geq \int_b^\infty b^2 f_X(x)dx = b^2 F^c_X(b) = b^2\beta. \tag{7}$$

(b) $\int_{-\infty}^{b-} f_X(x)dx = 1 - \int_b^\infty f_X(x)dx = 1 - \beta$. Similarly, since $\bar{X} = 0$,

$$\int_{-\infty}^{b-} x f_X(x)dx = 0 - \int_b^\infty x f_X(x)dx \leq -\int_b^\infty b f_X(x)dx = -b\beta. \tag{8}$$

(c) Let $h(x) = f_X(x)/(1-\beta)$ over $x \leq b$ and $h(x) = 0$ elsewhere. Then $h(x)$ is a PDF and the corresponding mean is $-b\beta/(1-\beta)$. The corresponding second moment is lower bounded by the square of that mean, so

$$\int_{-\infty}^{b-} x^2 h(x)dx \geq \frac{b^2\beta^2}{(1-\beta)^2}. \tag{9}$$

Since $h(x) = f_X(x)/(1-\beta)$, we have $\int_{-\infty}^{b-} x^2 f_X(x)dx \geq b^2\beta^2/(1-\beta)$.

(d) Substituting the results of (a) and (c) into (4)

$$\sigma^2 \geq \frac{b^2\beta^2}{1-\beta} + b^2\beta = \frac{b^2\beta}{1-\beta}. \tag{10}$$

This is met with equality when the inequalities in (a) and (c) are met with equality. Thus $X = b$ whenever $X \geq b$ and $X = -b\beta/(1-\beta)$ when $X < b$. Since β is the probability that $X \geq b$, we see that $p_X(b) = \beta$ and $p_X(-b\beta/(1-\beta)) = 1 - \beta$.

(e) From (10), $\sigma^2(1-\beta) \geq b^2\beta$, from which $\beta \leq \sigma^2/(b^2+\sigma^2)$, i.e., $\Pr\{X \geq b\} \leq \sigma^2/(b^2+\sigma^2)$. The conditions for equality are clearly the same as in (d). The final result follows by letting X be the fluctuation of Y and applying (3) to X.

3. Exercise 1.43 (Gallager's book) (**MS convergence → convergence in probability**) Assume that $\{Z_n : n \geq 1\}$ is a sequence of rv's and α is a number with the property that $\lim_{n\to\infty} \mathbb{E}[(Z_n-\alpha)^2] = 0$.

   (a) Let $\varepsilon > 0$ be arbitrary and show that for each $n \geq 0$,
$$\Pr\{|Z_n - \alpha| \geq \varepsilon\} \leq \frac{\mathbb{E}[(Z_n-\alpha)^2]}{\varepsilon^2}. \tag{11}$$

   (b) For the ε above, let $\delta > 0$ be arbitrary. Show that there is an integer m such that $\mathbb{E}[(Z_n-\alpha)^2] \leq \varepsilon^2\delta$ for all $n \geq m$.

   (c) Show that this implies convergence in probability.

   Solution:

   (a) It is a direct result of the Chebysev's inequality.

   (b) By the definition of a limit, $\lim_{n\to\infty} \mathbb{E}[(Z_n-\alpha)^2] = 0$ means that for all $\varepsilon_1 > 0$, there is an m large enough that $\mathbb{E}[(Z_n-\alpha)^2] < \varepsilon_1$ for all $n \geq m$. Choosing $\varepsilon_1 = \varepsilon^2\delta$ established the desired result.

   (c) Substituting $\mathbb{E}[(Z_n-\alpha)^2] \leq \varepsilon^2\delta$ into (11), we see that for all $\varepsilon, \delta > 0$, there is an m such that $\Pr\{|Z_n-\alpha|^2 \geq \varepsilon\} \leq \delta$ for all $n \geq m$. This is convergence in probability.

4. Exercise 1.44 (Gallager's book) Let $X_1, X_2, \ldots$ be a sequence of IID rv's each with mean 0 and variance σ^2. Let $S_n = X_1 + X_2 + \cdots + X_n$ for all n and consider the random variable $S_n/\sigma\sqrt{n} - S_{2n}/\sigma\sqrt{2n}$. Find the limiting CDF for this sequence of rv's as $n \to \infty$. The point of this exercise is to see clearly that the CDF of $S_n/\sigma\sqrt{n} - S_{2n}/\sigma\sqrt{2n}$ is converging in n but that the sequence of rv's is not converging in any reasonable sense.

   Solution: If we write out the above expression in terms of the X_i, we get
$$\frac{S_n}{\sigma\sqrt{n}} - \frac{S_{2n}}{\sigma\sqrt{2n}} = \sum_{i=1}^{n} X_i \left[\frac{1}{\sigma\sqrt{n}} - \frac{1}{\sigma\sqrt{2n}}\right] - \sum_{i=n+1}^{2n} \frac{X_i}{\sigma\sqrt{2n}} \tag{12}$$
$$= \frac{1-1/\sqrt{2}}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i - \frac{1}{\sigma\sqrt{2n}} \sum_{i=n+1}^{2n} X_i. \tag{13}$$

   By the central limit theorem, the first sum above approaches a zero-mean Gaussian distribution of variance $(1-1/\sqrt{2})^2$ and the second sum approaches a zero-mean Gaussian distribution of variance $1/2$. Since these two terms are independent, the difference approaches a Gaussian distribution with zero-mean and with variance $1/2 + (1-1/\sqrt{2})^2$.

   Note to students: I am not sure what the question means "the sequence of rv's is not converging in any reasonable sense". Let's ignore this part.

5. Exercise 1.48 (Gallager's book) Let $\{Y_n : n \geq 1\}$ be a sequence of rv's and assume that $\lim_{n\to\infty} \mathbb{E}[|Y_n|] = 0$. Show that $\{Y_n : n \geq 1\}$ converges to 0 in probability.

   Solution: Let $\epsilon > 0$ be fixed. Consider
$$\Pr(|Y_n - 0| > \epsilon) \leq \frac{\mathbb{E}[|Y_n|]}{\epsilon}$$

   Since $\lim_{n\to\infty} \mathbb{E}[|Y_n|] = 0$, we have that
$$\lim_{n\to\infty} \Pr(|Y_n - 0| > \epsilon) = 0$$

   which means that $Y_n \to 0$ in probability.

6. [Convergence of RVs] Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli(1/2). Define $Y_n = 2^n \prod_{i=1}^n X_i$. Does Y_n converge to 0 almost surely (with probability 1)? Does Y_n converge to 0 in mean square?

   Solution: Y_n converges to 0 a.s. Fix $m \in \mathbb{N}$ and then fix $0 < \epsilon < 2^m$. Then for this fixed m, and consider

$$\Pr(|Y_n - 0| < \epsilon, \forall n \geq m) = \Pr(X_n = 0 \text{ for some } n \leq m) = 1 - \Pr(X_n = 1, \forall n \leq m) = 1 - (1/2)^m$$

   which converges to 1 as $m \to \infty$. Hence $Y_n \to 0$ a.s.

   Y_n does not converge to 0 in m.s. Since

$$\mathbb{E}[(Y_n - 0)^2] = (1/2)^n 2^{2n} = 2^n \to \infty,$$

   the sequence doesn't converge in m.s.

7. [Two Independent Random Variables]

   Let X and Y be independent random variables, uniformly distributed on $[0, 2]$.

   (a) Find the mean and variance of XY.

   (b) Calculate the probability $\Pr(XY \leq 1)$.

   You may use, without proof, the fact that for a uniform random variable on $[a, b]$, the variance is $(b - a)^2/12$.

   Solution:

   (a) The mean is

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 1 \times 1 = 1$$

   Consider the second moment:

$$\mathbb{E}[(XY)^2] = \mathbb{E}[X^2]\mathbb{E}[Y^2] = [\mathsf{Var}(X) + (\mathbb{E}[X])^2][\mathsf{Var}(Y) + (\mathbb{E}[Y])^2] = (4/3)^2 = 16/9.$$

   Hence the variance is

$$\mathsf{Var}(XY) = \mathbb{E}[(XY)^2] - (\mathbb{E}[XY])^2 = 16/9 - 1^2 = 7/9.$$

   (b) We have

$$\Pr(XY \leq 1) = \Pr(Y \leq 1/X) \overset{(a)}{=} \int_0^2 f_X(x) \Pr(Y \leq 1/x)\, \mathrm{d}x = \int_0^2 F_Y(1/x) f_X(x)\, \mathrm{d}x,$$

   where (a) holds because X and Y are independent. Now observe that when $X \leq 1/2$, then $Y \leq 1/X$ with probability 1; and when $X \in (1/2, 2)$, then $Y \leq 1/X$ with probability $(1/X)/2$. Thus,

$$\Pr(XY \leq 1) = \int_0^{1/2} 1 \cdot \frac{1}{2}\, \mathrm{d}x + \int_{1/2}^2 \frac{1}{2x} \cdot \frac{1}{2}\, \mathrm{d}x = \frac{1}{4} + \frac{\log 2}{2}.$$

8. [Laws of Large Numbers]

   (a) State Chebychev's Inequality.

   (b) Suppose $\{X_i\}_{i=1}^\infty$ is a sequence of uncorrelated random variables (i.e., $\mathbb{E}[X_i X_j] = \mathbb{E}[X_i]\mathbb{E}[X_j]$ for $i \neq j$), each of which has finite mean, and assume that for all n, $\mathrm{Var}(X_n) \leq M < \infty$. Define

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i].$$

   and suppose that $\mu_n \to \mu$ as $n \to \infty$ and $|\mu| < \infty$. Let $S_n = \sum_{i=1}^n X_i$. Show that $S_n/n \to \mu$ in probability.

Solution:

(a) Suppose that the random variable X has finite mean μ. Then for any $c > 0$,

$$\Pr(|X - \mu| > c) \leq \frac{\mathsf{Var}(X)}{c^2}.$$

(b) Choose any $\epsilon > 0$. There exists N such that $|\mu_n - \mu| < \epsilon/2$ for all $n > N$. Thus for such $n > N$,

$$\begin{aligned}
\Pr\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) &\leq \Pr\left(\left|\frac{S_n}{n} - \mu_n\right| + |\mu_n - \mu| > \epsilon\right) \\
&\leq \Pr\left(\left|\frac{S_n}{n} - \mu_n\right| > \epsilon/2\right) \\
&\leq \frac{4\,\mathsf{Var}(S_n/n)}{\epsilon^2} \leq \frac{4M}{n\epsilon^2}
\end{aligned}$$

Thus for any $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) = 0$$

which implies $S_n/n \xrightarrow{\text{p}} \mu$.

9. [Convergence in Probability Implies Convergence in Distribution]

Show if $\{X_n\}$ convergence in distribution to X, then $\{X_n\}$ converges in distribution to the same rv X.

Hint: First convince yourself that

$$F_{X_n}(x) \leq F_X(x+\epsilon) + \Pr(|X_n - X| > \epsilon)$$

Consider the "flipped" inequality and use the definition of convergence in probability.

Solution: First we note that

$$\{X_n \leq x\} \subset \{X \leq x + \epsilon\} \cup \{|X_n - X| > \epsilon\}.$$

Why? Otherwise, if $X > x + \epsilon$ and $|X_n - X| \leq \epsilon$ (equivalently $X - \epsilon \leq X_n \leq X + \epsilon$), then $X_n > x$. By the union bound, we have

$$F_{X_n}(x) \leq F_X(x+\epsilon) + \Pr(|X_n - X| > \epsilon).$$

In the same way, one can show that

$$F_X(x-\epsilon) - \Pr(|X_n - X| > \epsilon) \leq F_{X_n}(x).$$

From convergence in probability, we know that for any $\epsilon > 0$, $\Pr(|X_n - X| > \epsilon) \to 0$. Thus, we have

$$F_X(x-\epsilon) \leq \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x+\epsilon),$$

and this statement holds for all $\epsilon > 0$. If x is a point of continuity of F_X, then $\lim_{\epsilon\to 0} F_X(x+\epsilon) = F_X(x)$. Thus, taking $\epsilon \to 0$ in the above display, and using the squeeze theorem (to be really pedantic), we get

$$F_X(x) = \lim_{n\to\infty} F_{X_n}(x)$$

for all points of continuity x of F_X, showing convergence in distribution.

10. [Probability Generating Function]

If X is a non-negative integer-valued rv then the function $Q(z)$ defined for $|z| \le 1$ by

$$Q(z) = \mathbb{E}[z^X] = \sum_{j=0}^{\infty} z^j \Pr(X = j)$$

is called the probability generating function of X.

In this problem you may interchange the differentiation operation and the infinite sum operation. The reason we can do this (for the more mathematically inclined students) is due to the so-called differentiable limit theorem and Abel's theorem.

(a) Show that

$$\frac{d^k}{dz^k} Q(z)\bigg|_{z=0} = k! \Pr(X = k).$$

(b) With 0 considered even, show that

$$\Pr(X \text{ is even}) = \frac{Q(-1) + Q(1)}{2}.$$

(c) If X is binomial with parameters n and p, show that

$$\Pr(X \text{ is even}) = \frac{1 + (1 - 2p)^n}{2}.$$

(d) If X is Poisson with mean λ, show that

$$\Pr(X \text{ is even}) = \frac{1 + e^{-2\lambda}}{2}.$$

Solutions:

(a) Consider the first derivative

$$\frac{d}{dz} \sum_{j=0}^{\infty} z^j p_X(j) = \sum_{j=0}^{\infty} \frac{d}{dz}(z^j p_X(j)) = \sum_{j=1}^{\infty} j z^{j-1} p_X(j)$$

Considering the k^{th} derivative, we get

$$\frac{d^k}{dz^k} \sum_{j=0}^{\infty} z^j p_X(j) = \sum_{j=k}^{\infty} j(j-1)\dots(j-k+1) z^{j-k} p_X(j)$$

Plugging $z = 0$ into the above display yields

$$\frac{d^k}{dz^k} \sum_{j=0}^{\infty} z^j p_X(j)\bigg|_{z=0} = k(k-1)\dots 1 \cdot p_X(k) = k! \Pr(X = k).$$

(b) We consider

$$Q(-1) + Q(1) = \sum_{j=0}^{\infty}((-1)^j p_X(j) + p_X(j)) = 2 \sum_{k=0}^{\infty} p_X(2k) = 2 \Pr(X \text{ is even})$$

as desired.

(c) For a binomial distribution,

$$Q(z) = \sum_{j=0}^{n} z^j \binom{n}{j} p^j (1-p)^{n-j} = \sum_{j=0}^{n} \binom{n}{j} (pz)^j (1-p)^{n-j} = [pz + (1-p)]^n$$

Thus, $Q(1) = 1$ and $Q(-1) = (1-2p)^n$, demonstrating the desired result upon using part (b).

(d) For a Poisson distribution,

$$Q(z) = \sum_{j=0}^{\infty} z^j \frac{e^{-\lambda}\lambda^j}{j!} = \sum_{j=0}^{\infty} \frac{e^{-\lambda}(z\lambda)^j}{j!} = e^{-\lambda+\lambda z} \sum_{j=0}^{\infty} \frac{e^{-\lambda z}(z\lambda)^j}{j!} = e^{\lambda(z-1)}.$$

Thus, $Q(1) = 1$ and $Q(-1) = e^{-2\lambda}$, demonstrating the desired result upon using part (b).