ECS 315: Probability and Random Processes **2014/1**

HW Solution 13 — Due: Not Due

Lecturer: Prapun Suksompong, Ph.D.

Problem 1. Complete the table below. Make sure that you use the definitions/notations that are presented in class.

$X \sim$	Support (S_X)	pmf/pdf on S_X	$\mathbb{E}X$	$\operatorname{Var} X$
$\mathcal{U}(\{n, n+1, \dots, n+d\})$				
Bernoulli(p)	$\{0,1\}$	$\begin{cases} 1-p, & x=0, \\ p, & x=1. \end{cases}$	p	$p(1-p)$
Binomial(n,p)				
$\mathcal{G}_0(\beta)$	$\{0,1,2,\dots\}$			
$\mathcal{G}_1(\beta)$	$\mathbb{N}$		$\frac{1}{1-\beta}$	
$\mathcal{P}(\alpha)$				
$\mathcal{U}(a,b)$				
$\mathcal{E}(\lambda)$		$\lambda e^{-\lambda x}$		
$\mathcal{N}(m,\sigma^2)$	$(-\infty,+\infty)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$	m	σ^2

Solution: Try to complete the table on your own. This table may be useful for the final exam.

Problem 2. In this question, we will explore the relationship between exponential random variable and geometric random variable.

(a) Start with an exponential random variable X whose parameter is λ. What is its pdf?

(b) What is the probability that X is in the interval $[a,b)$ for constants $b > a \geq 0$?

(c) What is the probability that X is in the interval $I_k = [(k-1)T, kT)$? Assume T is a positive real number and k is a positive integer. We will denote this probability by $p_k = P[X \in I_k]$.

(d) Consider the sequence of numbers $p_1, p_2, p_3, \dots$ where $p_k = P[X \in I_k]$ defined above. Are these p_k's agrees with a pmf of a geometric random variable? If so, what is the value of the parameter β for this geometric random variable?

(e) Let $Y = \lfloor X \rfloor$ where $\lfloor \cdot \rfloor$ is the floor function. Describe the random variable Y. (Continuous or discrete? What is its pdf/pmf?)

Solution:

(a) $f_X(x) = \boxed{\lambda e^{-\lambda x} u(x)}$ where $u(x) = 1_{[0,\infty)}(x)$.

(b) Recall that $F_X(x) = 1 - e^{\lambda x}$. So, $P[a \le X < b] = F(b) - F(a) = \boxed{e^{-a\lambda} - e^{-b\lambda}}$.

(c) $p_k = P[(k-1)T \le X < kT] = e^{-(k-1)T\lambda} - e^{-kT\lambda} = \boxed{e^{-(k-1)T\lambda}\left(1 - e^{-T\lambda}\right)}$. Note that this is valid for $k = 1, 2, \ldots$.

(d) Because $p_k = e^{-(k-1)T\lambda}\left(1 - e^{-T\lambda}\right) = \beta^{k-1}(1-\beta)$ where $\beta = e^{-T\lambda}$, the sequence p_k $\boxed{\text{agrees}}$ with the geometric$_1$ pmf with parameter $\boxed{\beta = e^{-T\lambda}}$.

(e) Here, $Y = k$ when $k \le X < k+1$. Note that Y is integer-valued and hence it is a $\boxed{\text{discrete}}$ random variable. At this point, you may try to redo all the calculations above to find $p_Y(k)$ which will be similar to what we have done to find p_k in the previous parts. Alternatively, we can try to use the old results by realizing that $p_Y(k) = p_{k+1}$ with $T = 1$. Hence, the pmf of Y is

$$p_Y(k) = \beta^k (1-\beta) \text{ where } \beta = e^{-T\lambda} \text{ for } k = 0, 1, 2, \ldots.$$

So, Y is $\boxed{\mathcal{G}_0(\beta) \text{ with } \beta = e^{-\lambda}}$.

See also Theorem 3.9 in [Yates & Goodman, 2005].

Problem 3. Consider the function

$$g(x) = \begin{cases} x, & x \ge 0 \\ 0, & x < 0. \end{cases}$$

The function g operates like a **half-wave rectifier** in that if a positive voltage x is applied, the output is $y = x$, while if a negative voltage x is applied, the output is $y = 0$. Suppose $Y = g(X)$, where $X \sim \mathcal{U}(-1, 1)$. Plot the cdf of Y.

Solution: See the handwritten solution. [Gubner, 2006, p. 197–198]

Problem 4. Suppose the elements of a $1 \times n$ random vector $\mathbf{A}$ and an $n \times n$ random matrix $\mathbf{B}$ are all i.i.d. random variables with shared pmf

$$p_X(x) = \begin{cases} 1/2, & x = 1, \\ 1/4, & x = 2, 3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Let $C = \frac{1}{n}\left(\mathbf{A} \times \mathbf{A}^T\right)$. Use the law of large numbers to estimate the value of C when n is large.

(b) Let $\mathbf{D} = \frac{1}{n}\left(\mathbf{B} \times \mathbf{B}^T\right)$. Use the law of large numbers to estimate the values of the elements inside matrix $\mathbf{D}$ when n is large.

Solution: First, note that

$$\begin{aligned} \mathbb{E}X &= \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 3 = \frac{7}{4} = 1.75, \text{ and} \\ \mathbb{E}\left[X^2\right] &= \frac{1}{2} \times 1^2 + \frac{1}{4} \times 2^2 + \frac{1}{4} \times 3^2 = \frac{15}{4} = 3.75. \end{aligned}$$

(a) Let A_i be the ith element in the random (row) vector $\mathbf{A}$. Note that $C = \frac{1}{n}\left(\mathbf{A} \times \mathbf{A}^T\right) = \frac{1}{n}\sum_{i=1}^{n} A_i^2$. Now, because the A_i's are i.i.d., we also know that the A_i^2's are i.i.d. By the law of large numbers, this implies

$$C \to \mathbb{E}\left[A_i^2\right] = \mathbb{E}\left[X^2\right] = 3.75.$$

Therefore, when n is large, C will be $\boxed{\text{close to } 3.75}$.

(b) For a matrix $\mathbf{M}$, define $M_{i,j}$ to be the entry in the ith row and jth column of the matrix $\mathbf{M}$. Now,

$$D_{i,j} = \frac{1}{n}\sum_{k=1}^{n} B_{i,k}\left(B^T\right)_{k,j} = \frac{1}{n}\sum_{k=1}^{n} B_{i,k}B_{j,k}.$$

Let $Y_k = B_{i,k}B_{j,k}$. Then, note that $Y_1, Y_2, \ldots, Y_n$ are i.i.d. Therefore, as $n \to \infty$,

$$D_{i,j} \to E\left[B_{i,k}B_{j,k}\right] = \begin{cases} \mathbb{E}\left[B_{i,k}^2\right], & i = j, \\ \mathbb{E}\left[B_{i,k}\right]\mathbb{E}\left[B_{j,k}\right], & i \neq j, \end{cases} = \begin{cases} \mathbb{E}\left[X^2\right], & i = j, \\ \left(\mathbb{E}\left[X\right]\right)^2, & i \neq j. \end{cases}$$

Therefore, when n is large,

- the (main) **diagonal entries** of matrix $\mathbf{D}$ will be $\boxed{\text{close to } 3.75}$ and

- the **entries outside the main diagonal** will be $\boxed{\text{close to } 1.75^2 = 3.0625}$.

Problem 5. A student has passed a final exam by supplying correct answers for 26 out of 50 multiple-choice questions. For each question, there was a choice of three possible answers, of which only one was correct. The student claims not to have learned anything in the course and not to have studied for the exam, and says that his correct answers are the product of guesswork. Use Table 3.1 and/or Table 3.2 from [Yates & Goodman, 2005] to determine whether you should believe him.

Solution: This problem can be approached as follows: take as hypothesis that the student did guess at all the answers and calculate the probability of identifying 26 or more correct answers through guesswork. If this probability is below a threshold value you have chosen in advance, you judge that the student is bluffing.

If all the answers are guessed at, then the number of correct answers can be seen as the number of successes in $n = 50$ independent trials of a Bernoulli experiment having a success probability of $p = 1/3$. The binomial probability model is thus applicable.

A generally useful method of determining whether 26 correct answers is exceptional is based on finding out how many standard deviations lie between the observed number of correct answers achieved and the expected number. To do so, a quick approach is to approximate the binomial distribution with parameters $n = 50$ and $p = 1/3$ by a normal distribution with expected value $np = 50/3 \approx 16.67$ and standard deviation $\sqrt{np(1-p)} = 10/3 \approx 3.33$. So, the observed value of 26 correct answers lies

$$\frac{26 - \frac{50}{3}}{\frac{10}{3}} = 2.8$$

standard deviations above the expected value.

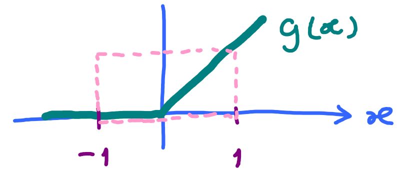
Note that it is useful to remember as a rule of thumb that the probability of a normally distributed random variable taking on a value lying three or more standard deviations above the expected value is very small (the probability is 0.00135). To be more precise, from Table 3.1, we get $1 - \Phi(2.8) \approx 1 - 0.99744 = 0.0026$.

From the analysis above, we see that the probability of such a deviation occurring is quite small. There is very good reason, therefore, to suppose that the student is bluffing, and that he in fact did prepare for the exam.

HW13 Q3

Friday, October 12, 2012 4:33 PM

As X varies over $[-1,1]$, $Y=g(X)$ varies over $[0,1]$.



This immediately gives us $F_Y(y) = \begin{cases} 0, & y<0, \\ 1, & y \geq 1. \end{cases}$

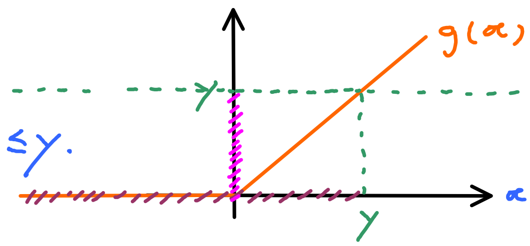
Useful fact : If you know that your random variable always $\in [a,b]$, then $F_Y(y) = \begin{cases} 0, & y<a \\ 1, & y \geq b. \end{cases}$

Of course, we still need to find $F_Y(y)$ for $0 \leq y < 1$.

To do this, reconsider the function g.

Observe that for $0 \leq y < 1$,

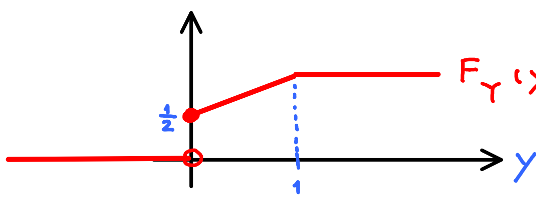
$g(x) \leq y$ if and only if $x \leq y$.



Therefore,

$$F_Y(y) = P[Y \leq y] = P[X \leq y] = F_X(y) = \underbrace{\frac{y-(-1)}{1-(-1)}}_{\text{formular for cdf of uniform r.v.}} = \frac{y+1}{2}$$

So,



$F_Y(y) = \begin{cases} 0, & y<0 \\ (y+1)/2, & 0 \leq y < 1 \\ 1, & y \geq 1. \end{cases}$