ECS 315: Probability and Random Processes **2013/1**

HW Solution 9 — Due: Sep 13

Lecturer: Prapun Suksompong, Ph.D.

Instructions

(a) ONE part of a question will be graded (5 pt). Of course, you do not know which part will be selected; so you should work on all of them.

(b) It is important that you try to solve all problems. (5 pt)

(c) Late submission will be heavily penalized.

(d) Write down all the steps that you have done to obtain your answers. You may not get full credit even when your answer is correct without showing how you get your answer.

Problem 1 (Randomly Phased Sinusoid)**.** Suppose Θ is a uniform random variable on the interval $(0, 2\pi)$.

(a) Consider another random variable X defined by

$$X = 5\cos(7t + \Theta)$$

where t is some constant. Find $\mathbb{E}[X]$.

(b) Consider another random variable Y defined by

$$Y = 5\cos(7t_1 + \Theta) \times 5\cos(7t_2 + \Theta)$$

where t_1 and t_2 are some constants. Find $\mathbb{E}[Y]$.

Solution: First, because Θ is a uniform random variable on the interval $(0, 2\pi)$, we know that $f_\Theta(\theta) = \frac{1}{2\pi}1_{(0,2\pi)}(t)$. Therefore, for "any" function g, we have

$$\mathbb{E}[g(\Theta)] = \int_{-\infty}^{\infty} g(\theta) f_\Theta(\theta) d\theta.$$

(a) X is a function of Θ. $\mathbb{E}[X] = 5\mathbb{E}[\cos(7t + \Theta)] = 5\int_0^{2\pi} \frac{1}{2\pi}\cos(7t + \theta)d\theta$. Now, we know that integration over a cycle of a sinusoid gives 0. So, $\mathbb{E}[X] = \boxed{0}$.

(b) Y is another function of Θ.

$$\begin{aligned}\mathbb{E}[Y] &= \mathbb{E}[5\cos(7t_1+\Theta)\times 5\cos(7t_2+\Theta)] = \int_0^{2\pi}\frac{1}{2\pi}5\cos(7t_1+\theta)\times 5\cos(7t_2+\theta)d\theta \\ &= \frac{25}{2\pi}\int_0^{2\pi}\cos(7t_1+\theta)\times\cos(7t_2+\theta)d\theta.\end{aligned}$$

Recall[1] the cosine identity

$$\cos(a)\times\cos(b) = \frac{1}{2}\left(\cos(a+b)+\cos(a-b)\right).$$

Therefore,

$$\begin{aligned}\mathbb{E}Y &= \frac{25}{4\pi}\int_0^{2\pi}\cos(14t+2\theta)+\cos(7(t_1-t_2))\,d\theta \\ &= \frac{25}{4\pi}\left(\int_0^{2\pi}\cos(14t+2\theta)\,d\theta+\int_0^{2\pi}\cos(7(t_1-t_2))\,d\theta\right).\end{aligned}$$

The first integral gives 0 because it is an integration over two period of a sinusoid. The second integrand in the second integral is a constant. So,

$$\mathbb{E}Y = \frac{25}{4\pi}\cos(7(t_1-t_2))\int_0^{2\pi}d\theta = \frac{25}{4\pi}\cos(7(t_1-t_2))\,2\pi = \boxed{\frac{25}{2}\cos(7(t_1-t_2))}.$$

Problem 2 (Yates and Goodman, 2005, Q3.4.5)**.** X is a continuous uniform $(-5,5)$ random varible.

(a) What is its pdf $f_X(x)$?

(b) What is its cdf $F_X(x)$?

(c) What is $\mathbb{E}[X]$?

(d) What is $\mathbb{E}[X^5]$?

---

[1]This identity could be derived easily via the Euler's identity:

$$\begin{aligned}\cos(a)\times\cos(b) &= \frac{e^{ja}+e^{-ja}}{2}\times\frac{e^{jb}+e^{-jb}}{2} = \frac{1}{4}\left(e^{ja}e^{jb}+e^{-ja}e^{jb}+e^{ja}e^{-jb}+e^{-ja}e^{-jb}\right) \\ &= \frac{1}{2}\left(\frac{e^{ja}e^{jb}+e^{-ja}e^{-jb}}{2}+\frac{e^{-ja}e^{jb}+e^{ja}e^{-jb}}{2}\right) \\ &= \frac{1}{2}\left(\cos(a+b)+\cos(a-b)\right).\end{aligned}$$

(e) What is $\mathbb{E}\left[e^X\right]$?

Solution: For a uniform random variable X on the interval (a,b), we know that

$$f_X(x) = \begin{cases} 0, & x < a \text{ or } x > b, \\ \frac{1}{b-a}, & a \le x \le b \end{cases}$$

and

$$F_X(x) = \begin{cases} 0, & x < a \text{ or } x > b, \\ \frac{x-a}{b-a}, & a \le x \le b. \end{cases}$$

In this problem, we have $a = -5$ and $b = 5$.

(a) $f_X(x) = \boxed{\begin{cases} 0, & x < -5 \text{ or } x > 5, \\ \frac{1}{10}, & -5 \le x \le 5 \end{cases}}$

(b) $F_X(x) = \boxed{\begin{cases} 0, & x < a \text{ or } x > b, \\ \frac{x+5}{10}, & a \le x \le b. \end{cases}}$

(c) In general,

$$\mathbb{E}X = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}\left.\frac{x^2}{2}\right|_a^b = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2}.$$

With $a = -5$ and $b = 5$, we have $\mathbb{E}X = \boxed{0}$.

(d) In general,

$$\mathbb{E}\left[X^5\right] = \int_a^b x^5\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b x^5dx = \frac{1}{b-a}\left.\frac{x^6}{6}\right|_a^b = \frac{1}{b-a}\frac{b^6-a^6}{2}.$$

With $a = -5$ and $b = 5$, we have $\mathbb{E}\left[X^5\right] = \boxed{0}$.

(e) In general,

$$\mathbb{E}\left[e^X\right] = \int_a^b e^x\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b e^xdx = \frac{1}{b-a}\left.e^x\right|_a^b = \frac{e^b-e^a}{b-a}.$$

With $a = -5$ and $b = 5$, we have $\mathbb{E}\left[e^X\right] = \boxed{\frac{e^5-e^{-5}}{10}} \approx 14.84$.

Problem 3. A random variable X is a Gaussian random variable if its pdf is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2},$$

for some constants m and σ. Furthermore, when a Gaussian random variable has $m = 0$ and $\sigma = 1$, we say that it is a standard Gaussian random variable. There is no closed-form expression for the cdf of the standard Gaussian random variable. The cdf itself is denoted by Φ and its values (or its complementary values $Q(\cdot) = 1 - \Phi(\cdot)$) are traditionally provided by a table. We refer to this kind of table as the Φ table. Examples of such tables are Table 3.1 and Table 3.2 in [Y&G].

Suppose Z is a standard Gaussian random variable.

(a) Use the Φ table to find the following probabilities:

- (i) $P[Z < 1.52]$
- (ii) $P[Z < -1.52]$
- (iii) $P[Z > 1.52]$
- (iv) $P[Z > -1.52]$
- (v) $P[-1.36 < Z < 1.52]$

(b) Use the Φ table to find the value of c that satisfies each of the following relation.

- (i) $P[Z > c] = 0.14$
- (ii) $P[-c < Z < c] = 0.95$

Solution:

(a)

- (i) $P[Z < 1.52] = \Phi(1.52) = \boxed{0.9357.}$
- (ii) $P[Z < -1.52] = \Phi(-1.52) = 1 - \Phi(1.52) = 1 - 0.9357 = \boxed{0.0643.}$
- (iii) $P[Z > 1.52] = 1 - P[Z < 1.52] = 1 - 0.9357 = \boxed{0.0643.}$
- (iv) It is straightforward to see that the area of $P[Z > -1.52]$ is the same as $P[Z < 1.52] = \Phi(1.52)$. So, $P[Z > -1.52] = \boxed{0.9357.}$
  Alternatively, $P[Z > -1.52] = 1 - P[Z \le -1.52] = 1 - \Phi(-1.52) = 1 - (1 - \Phi(1.52)) = \Phi(1.52)$.
- (v) $P[-1.36 < Z < 1.52] = \Phi(1.52) - \Phi(-1.36) = \Phi(1.52) - (1 - \Phi(1.36)) = \Phi(1.52) + \Phi(1.36) - 1 = 0.9357 + 0.9131 - 1 = \boxed{0.8488.}$

(b)

(i) $P[Z > c] = 1 - P[Z \le c] = 1 - \Phi(c)$. So, we need $1 - \Phi(c) = 0.14$ or $\Phi(c) = 1 - 0.14 = 0.86$. In the Φ table, we do not have exactly 0.86, but we have 0.8599 and 0.8621. Because 0.86 is closer to 0.8599, we answer the value of c whose $\phi(c) = 0.8599$. Therefore, $c \approx \boxed{1.08.}$

(ii) $P[-c < Z < c] = \Phi(c) - \Phi(-c) = \Phi(c) - (1 - \Phi(c)) = 2\Phi(c) - 1$. So, we need $2\Phi(c) - 1 = 0.95$ or $\Phi(c) = 0.975$. From the Φ table, we have $c \approx \boxed{1.96.}$

Problem 4. The peak temperature T, as measured in degrees Fahrenheit, on a July day in New Jersey is a $\mathcal{N}(85, 100)$ random variable.

Remark: Do not forget that, for our class, the second parameter in $\mathcal{N}(\cdot, \cdot)$ is the variance (not the standard deviation).

(a) Express the cdf of T in terms of the Φ function

Hint: Recall that the cdf of a random variable T is given by $F_T(t) = P[T \le t]$. For $T \sim \mathcal{N}(m, \sigma^2)$,

$$F_T(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx.$$

The Φ function, which is the cdf of the standard Gaussian random variable, is given by

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx.$$

(b) Express each of the following probabilities in terms of the Φ function(s). Make sure that the arguments of the Φ functions are positive. (Positivity is required so that we can directly use the Φ/Q tables to evaluate the probabilities.)

(i) $P[T > 100]$

(ii) $P[T < 60]$

(iii) $P[70 \le T \le 100]$

(c) Express each of the probabilities in part (b) in terms of the Q function(s). Again, make sure that the arguments of the Q functions are positive.

(d) Evaluate each of the probabilities in part (b) using the Φ/Q tables.

(e) Observe that Table 3.1 stops at $z = 2.99$ and Table 3.2 starts at $z = 3.00$. Why is it better to give a table for $Q(z)$ instead of $\Phi(z)$ when z is large?

Solution:

(a) Continue from the hint. We perform a change of variable using $y = \frac{x-m}{\sigma}$ to get

$$F_T(t) = \int_{-\infty}^{\frac{x-m}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy = \Phi\left(\frac{x-m}{\sigma}\right).$$

Here, $T \sim \mathcal{N}(85, 10^2)$. Therefore, $F_T(t) = \boxed{\Phi\left(\frac{t-85}{10}\right).}$

(b)

(i) $P[T > 100] = 1 - P[T \le 100] = 1 - F_T(100) = 1 - \Phi\left(\frac{100-85}{10}\right) = 1 - \Phi(1.5)$

(ii) $P[T < 60] = P[T \le 60]$ because T is a continuous random variable and hence $P[T = 60] = 0$. Now, $P[T \le 60] = F_T(60) = \Phi\left(\frac{60-85}{10}\right) = \Phi(-2.5) = \boxed{1 - \Phi(2.5).}$ Note that, for the last equality, we use the fact that $\Phi(-x) = 1 - \Phi(x)$.

(iii)

$$\begin{aligned} P[70 \le T \le 100] &= F_T(100) - F_T(70) = \Phi\left(\frac{100-85}{10}\right) - \Phi\left(\frac{70-85}{10}\right) \\ &= \Phi(1.5) - \Phi(-1.5) = \Phi(1.5) - (1 - \Phi(1.5)) = \boxed{2\Phi(1.5) - 1.} \end{aligned}$$

(c) In this question, we use the fact that $Q(x) = 1 - \Phi(x)$.

(i) $1 - \Phi(1.5) = \boxed{Q(1.5).}$

(ii) $1 - \Phi(2.5) = \boxed{Q(2.5).}$

(iii) $2\Phi(1.5) - 1 = 2(1 - Q(1.5)) - 1 = 2 - 2Q(1.5) - 1 = \boxed{1 - 2Q(1.5).}$

(d)

(i) $1 - \Phi(1.5) = 1 - 0.9332 = \boxed{0.0668.}$

(ii) $1 - \Phi(2.5) = 1 - 0.99379 = \boxed{0.0062.}$

(iii) $2\Phi(1.5) - 1 = 2(0.9332) - 1 = \boxed{0.8664.}$

(e) When z is large, $\Phi(z)$ will start with 0.999... The first few significant digits will all be the same and hence not quite useful to be there.

Problem 5. (Function of Continuous Random Variable) Let $X \sim \mathcal{E}(5)$ and $Y = 2/X$. Find

(a) $F_Y(y)$.

(b) $f_Y(y)$.

(c) $\mathbb{E}Y$

Hint: Because $\frac{d}{dy}e^{-\frac{10}{y}} = \frac{10}{y^2}e^{-\frac{10}{y}} > 0$ for $y \neq 0$. We know that $e^{-\frac{10}{y}}$ is an increasing function on our range of integration. In particular, consider $y > 10/\ln(2)$. Then, $e^{-\frac{10}{y}} > \frac{1}{2}$. Hence,

$$\int_0^\infty \frac{10}{y}e^{-\frac{10}{y}}dy > \int_{10/\ln 2}^\infty \frac{10}{y}e^{-\frac{10}{y}}dy > \int_{10/\ln 2}^\infty \frac{10}{y}\frac{1}{2}dy = \int_{10/\ln 2}^\infty \frac{5}{y}dy.$$

Remark: To be technically correct, we should be a little more careful when writing $Y = \frac{2}{X}$ because it is undefined when $X = 0$. Of course, this happens with 0 probability; so it won't create any serious problem. However, to avoid the problem, we may define Y by

$$Y = \begin{cases} 2/X, & X \neq 0, \\ 0, & X = 0. \end{cases}$$

Solution:

(a) Because $X > 0$, we know that $Y > 0$ and hence, $F_Y(y) = 0$ for $y \leq 0$. Note that Y can not be $= 0$. We need $X = \infty$ or $-\infty$ to make $Y = 0$. However, $\pm\infty$ are not real numbers therefore they are not possible X values.

For $y > 0$,

$$F_Y(y) = P[Y \leq y] = P\left[\frac{2}{X} \leq y\right] = P\left[X \geq \frac{2}{y}\right] = 1 - F_X\left(\frac{2}{y}\right)$$
$$= e^{-5\left(\frac{2}{y}\right)} = e^{\frac{-10}{y}}$$

Hence,

$$F_Y(y) = \boxed{\begin{cases} e^{-\frac{10}{y}}, & y > 0 \\ 0, & y \leq 0 \end{cases}}$$

(b) Because we have already derive the cdf in the previous part, we can find the pdf via the cdf by $f_Y(y) = \frac{d}{dy}F_Y(y)$. This gives f_Y at all points except at $y = 0$ which we will set f_Y to be 0 there. Hence,

$$f_Y(y) = \boxed{\begin{cases} \frac{10}{y^2}e^{-\frac{10}{y}}, & y > 0 \\ 0, & y \le 0. \end{cases}}$$

Remark: In general, for $Y = \frac{a}{X}$, $f_Y(y) = \left|\frac{a}{y^2}\right| f_X\left(\frac{a}{y}\right)$.

(c) From the hint, we have

$$\begin{aligned}
\mathbb{E}Y &= \int_0^\infty y\frac{10}{y^2}e^{-\frac{10}{y}}\,dy = \int_0^\infty \frac{10}{y}e^{-\frac{10}{y}}\,dy \\
&> \int_{10/\ln 2}^\infty \frac{10}{y}e^{-\frac{10}{y}}\,dy > \int_{10/\ln 2}^\infty \frac{10}{y}\frac{1}{2}dy = \int_{10/\ln 2}^\infty \frac{5}{y}dy \\
&= 5\ln y\big|_{10/\ln 2}^{\infty} = \infty.
\end{aligned}$$

Therefore, $\mathbb{E}Y = \boxed{\infty}$.

Extra Questions

Here are some questions for those who want extra practice.

Problem 6. Let X be a uniform random variable on the interval $[0,1]$. Set

$$A = \left[0, \frac{1}{2}\right), \quad B = \left[0, \frac{1}{4}\right) \cup \left[\frac{1}{2}, \frac{3}{4}\right), \quad \text{and } C = \left[0, \frac{1}{8}\right) \cup \left[\frac{1}{4}, \frac{3}{8}\right) \cup \left[\frac{1}{2}, \frac{5}{8}\right) \cup \left[\frac{3}{4}, \frac{7}{8}\right).$$

Are the events $[X \in A], [X \in B]$, and $[X \in C]$ independent?

Solution: Note that

$$P[X \in A] = \int_0^{\frac{1}{2}} dx = \frac{1}{2},$$

$$P[X \in B] = \int_0^{\frac{1}{4}} dx + \int_{\frac{1}{2}}^{\frac{3}{4}} dx = \frac{1}{2}, \text{ and}$$

$$P[X \in C] = \int_0^{\frac{1}{8}} dx + \int_{\frac{1}{4}}^{\frac{3}{8}} dx + \int_{\frac{1}{2}}^{\frac{5}{8}} dx + \int_{\frac{3}{4}}^{\frac{7}{8}} dx = \frac{1}{2}.$$

Now, for pairs of events, we have

$$P([X \in A] \cap [X \in B]) = \int_0^{\frac{1}{4}} dx = \frac{1}{4} = P[X \in A] \times P[X \in B], \tag{9.1}$$

$$P([X \in A] \cap [X \in C]) = \int_0^{\frac{1}{8}} dx + \int_{\frac{1}{4}}^{\frac{3}{8}} dx = \frac{1}{4} = P[X \in A] \times P[X \in C], \text{ and} \tag{9.2}$$

$$P([X \in B] \cap [X \in C]) = \int_0^{\frac{1}{8}} dx + \int_{\frac{1}{2}}^{\frac{5}{8}} dx = \frac{1}{4} = P[X \in B] \times P[X \in C]. \tag{9.3}$$

Finally,

$$P([X \in A] \cap [X \in B] \cap [X \in C]) = \int_0^{\frac{1}{8}} dx = \frac{1}{8} = P[X \in A]\, P[X \in B]\, P[X \in C]. \tag{9.4}$$

From (9.1), (9.2), (9.3) and (9.4), we can conclude that the events $[X \in A], [X \in B]$, and $[X \in C]$ are $\boxed{\text{independent}}$.

Problem 7 (Q3.5.6)**.** Solve this question using Table 3.1 and/or Table 3.2 from [Yates & Goodman, 2005]:

A professor pays 25 cents for each blackboard error made in lecture to the student who points out the error. In a career of n years filled with blackboard errors, the total amount in dollars paid can be approximated by a Gaussian random variable Y_n with expected value $40n$ and variance $100n$.

(a) What is the probability that Y_{20} exceeds 1000?

(b) How many years n must the professor teach in order that $P[Y_n > 1000] > 0.99$?

Solution: We are given[2] that $Y_n \sim \mathcal{N}(40n, 100n)$. Recall that when $X \sim \mathcal{N}(m, \sigma^2)$,

$$F_X(x) = \Phi\left(\frac{x-m}{\sigma}\right). \tag{9.5}$$

(a) Here $n = 20$. So, we have $Y_n \sim \mathcal{N}(40 \times 20, 100 \times 20) = \mathcal{N}(800, 2000)$. For this random variable $m = 800$ and $\sigma = \sqrt{2000}$.

We want to find $P[Y_{20} > 1000]$ which is the same as $1 - P[Y_20 \le 1000]$. Expressing this quantity using cdf, we have

$$P[Y_{20} > 1000] = 1 - F_{Y_{20}}(1000).$$

Apply (9.5) to get

$$P[Y_{20} > 1000] = 1 - \Phi\left(\frac{1000-800}{\sqrt{2000}}\right) = 1 - \Phi(4.472) \approx Q(4.47) \approx \boxed{3.91 \times 10^{-6}}.$$

(b) Here, the value of n is what we want. So, we will need to keep the formula in the general form. Again, from (9.5), for $Y_n \sim \mathcal{N}(40n, 100n)$, we have

$$P[Y_n > 1000] = 1 - F_{Y_n}(1000) = 1 - \Phi\left(\frac{1000-40n}{10\sqrt{n}}\right) = 1 - \Phi\left(\frac{100-4n}{\sqrt{n}}\right).$$

[2]Note that the expected value and the variance in this question are proportional to n. This naturally occurs when we consider the sum of i.i.d. random variables. The approximation by Gaussian random variable is a result of the central limit theorem (CLT).

To find the value of n such that $P[Y_n > 1000] > 0.99$, we will first find the value of z which make

$$1 - \Phi(z) > 0.99. \tag{9.6}$$

At this point, we may try to solve for the value of Z by noting that (9.6) is the same as

$$\Phi(z) < 0.01. \tag{9.7}$$

Unfortunately, the tables that we have start with $\Phi(0) = 0.5$ and increase to something close to 1 when the argument of the Φ function is large. This means we can't directly find 0.01 in the table. Of course, 0.99 is in there and therefore we will need to solve (9.6) via another approach.

To do this, we use another property of the Φ function. Recall that $1 - \Phi(z) = \Phi(-z)$. Therefore, (9.6) is the same as

$$\Phi(-z) > 0.99. \tag{9.8}$$

From our table, we can then conclude that (9.7) (which is the same as (9.8)) will happen when $-z > 2.33$. (If you have `MATLAB`, then you can get a more accurate answer of 2.3263.)

Now, plugging in $z = \frac{100-4n}{\sqrt{n}}$, we have $\frac{4n-100}{\sqrt{n}} > 2.33$. To solve for n, we first let $x = \sqrt{n}$. In which case, we have $\frac{4x^2-100}{x} > 2.33$ or, equivalently, $4x^2 - 2.33x - 100 > 0$. The two roots are $x = -4.717$ and $x > 5.3$. So, We need $x < -4.717$ or $x > 5.3$. Note that $x = \sqrt{n}$ and therefore can not be negative. So, we only have one case; that is, we need $x > 5.3$. Because $n = x^2$, we then conclude that we need $\boxed{n > 28.1 \text{ years}}$.