Probability Theory for EOR 2021/2022

—

Solutions to Exercises

—

Week 5

December 17, 2021

PDFs and CDFs

2.a You friend tell you he/she would phone you tomorrow morning, but does not leave a valid time slot and could happen arbitrarily from 0 to 1/2 (normalize 7am to 12 am). $f(x) = 2, \forall x \in [0, 1/2]$ and $f(x) = 0$ otherwise.

Verify this is a PDF:

- non-negative: $f(x) \geq 0$
- integrates to 1: $\sum_{\mathbb{R}} f(x)dx = 1$

2.b Consider one case such that $f(x) \geq 1, \forall x \in [a, b]$ and $f(x) = 0$ otherwise. Then we know

$$1 = \int_a^b f(x)dx \geq \int_a^b 1dx = b - a$$

so the length of the interval is less than 1.

Uniform and universality of the Uniform

10.a Uniform probability is proportional to the interval length $(2 + 4)/8$.

Here we can calculate using the fact that for a continuous r.v. X with given PDF $f(.)$ we can simply integrate over the region A to calculate $\mathbb{P}(X \in A) = \int_A f(a)da$:

$$\mathbb{P}\left(U \in (0,2) \cup (3,7)\right) = \int_{(0,2)\cup(3,7)} 1/(8-0)da = \int_0^2 1/8da + \int_3^7 1/8da = 3/4$$

10.b Unif(3 , 7). (Which proposition in the book?)

Take a look of the proposition 5.2.3.

Conditional distribution is just one special case of conditional probability, where we look into the probability of a r.v. being smaller than a arbitrary constant (this part resembles the definition of CDF) conditional on certain events (take a look at page 46, the definition of conditional probability.) Now we derive the conditional distribution using the definition of the conditional probability: For an arbitrary constant $u \in (3,7)$

$$\begin{aligned}\mathbb{P}\left(U \leq u | U \in (3,7)\right) &= \mathbb{P}\left(\{U \leq u\} \cap \{U \in (3,7)\}\right) / \mathbb{P}\left(U \in (3,7)\right)\\ &= \mathbb{P}\left(U \in (3,u)\right) / \mathbb{P}\left(U \in (3,7)\right)\\ &= \left(\int_3^u 1/8da\right) / \left(\int_3^7 1/8da\right)\\ &= (u-3)/4\end{aligned}$$

To complete, for $u \leq 3, \mathbb{P}\left(U \leq u | U \in (3,7)\right) = 0$; for $u \geq 7, \mathbb{P}\left(U \leq u | U \in (3,7)\right) = 1$.

Note that $\mathbb{P}\left(U \leq u | U \in (3,7)\right)$ is a function of u, and this function is equal to the CDF of Unif(3,7)!

12 Denote the length of one piece as L, the other would be $1-L$, and $L\ \sim$Unif(0,1). This question is equivalent to deriving the distribution of $X = \max\{L, 1-L\}$. For $l \in [1/2, 1)$

$$\mathbb{P}(X \leq l) = \mathbb{P}(L \leq l, 1-L < l) = \mathbb{P}(1-l < L \leq l) = \int_{1-l}^{l} 1/(1-0)dl = 2l-1$$

where the second last equality we use the PDF of Unif(0,1). To complete, for $l \geq 1, \mathbb{P}(X \leq l) = 1$; $l < 1/2, \mathbb{P}(X \leq l) = 0$. The distribution we derive is the same as a distribution of a Unif(1/2,1) as the CDF is linearly increasing from 0 to 1 over the interval (1/2,1), and thus $X\ \sim$Unif(1/2,1), which then the properties of Uniform distribution tell us that $\mathbb{E}(X) = ((1/2)+1)/2 = 3/4$.

Alternatively, we can check the PDF, by taking derivatives of $\mathbb{P}(X \leq l)$, denote $f(l) = \frac{\partial \mathbb{P}(X \leq l)}{\partial l}(l)$, then we know

$$f(l) = \begin{cases} 2 & l \in (1/2, 1) \\ 0 & otherwise \end{cases} \tag{1}$$

which is the PDF of Unif(1/2,1). (technically, as a function of l, $\mathbb{P}(X \leq l)$ has kinks when $l = 0, l = 1$ and thus at these two points, the function is not differentiable. As one convention, we normally assign zero values **to its derivatives (PDF)** at these points. Technically speaking, you could assign any values you like **to its derivatives**

(PDF) at these non-differentiable points for a continuous r.v., as that will not change the integration value and thus will not change the probability.)[1]

What we can learn from this exercise is to learn the type of a certain r.v., we only need to look at their CDFs (MGF is also one choice, or PDF, but to derive PDF, normally it is easy to derive CDF first). When you calculating the CDF, do complete the whole function by giving the function values along the whole real line. Furthermore, we also revisit some math knowledge:

- *not differentiable does not imply not continuous, but not continuous does imply not differentiable*
- *What is more, is that you can have many PDFs (they can differ in at most countably many points) corresponding to the same distribution (CDF) and the PDFs of one continuous r.v. whose CDF is always continuous (but not necessarily differentiable at all points) can be discontinuous (can take arbitrary values at these points) at these non-differentiable points. By the properties of Rieman integration, you will notice these PDFs (they can differ in at these non-differentiable points (at most countably many points)) will lead to the same integration values and thus the same probability. As one convention in the textbook, to avoid such confusion, you will see the textbook defines all continuous r.v.s by one given PDFs (instead of CDFs, some other books use CDFs and would lead to many PDFs associated with one same continuous r.v.) and simply assigns zero values to the PDFs at these non-differentiable points (take the Definition 5.2.1 for example, look at the PDF values at a, and b).*

16.a By **LOTUS** and the PDF of Unif(0,1) takes the value 1 within the interval (0,1) and 0 elsewhere, we have

$$\int_0^1 \left[\log\left(u/(1-u)\right)\right]^2 du$$

Further comment: Note that X^2 is also a function of U: $X^2 = (\log(U/(1-U)))^2$. So LOTUS allows us to derive $\mathbb{E}X^2 = \mathbb{E}(\log(U/(1-U)))^2$ Using the distribution of U as well.

Alternatively (this one takes more computation time and is slightly complicated), if you want to try, you can also derive X's PDF, which should be $f_X(x) = e^x/(e^x+1)^2$

[1]Thank Moesen Tajik, whose questions help to polish the comments

with support $(-\infty, \infty)$*. Then again by LOTUS*

$$\mathbb{E}X^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_{-\infty}^{\infty} x^2 e^x/(e^x+1)^2 dx$$

which is also a correct answer. We can use Integration by substitution to show these two integrals are equal. Choose $x = \log(u/(1-u))$ *and thus*

$$\int_{-\infty}^{\infty} x^2 e^x/(e^x+1)^2 dx = \int_0^1 (\log(u/(1-u)))2(u/(1-u))/(u/(1-u)+1)^2 d\log(u/(1-u))$$

which after simplifying would be $\int_0^1 [\log(u/(1-u))]^2 du$*. Both expressions* $\int_0^1 [\log(u/(1-u))]^2 du$ *and* $\int_{-\infty}^{\infty} x^2 e^x/(e^x+1)^2 dx$ *are correct answers (there are more different expressions that could all lead to the same integration value, and they could all be correct).*

16.b Use the fact that (why we have this, think about the hint given in the exercise) $\mathbb{E}X = \mathbb{E}(-X)$, we have

$$\mathbb{E}X = \mathbb{E}(X - X)/2 = 0$$

Here we give two ways of proving the result (all rely on the fact that if two **two r.v.s have the same distribution then they have the same expectation value**) :

(1) First approach (I explain the above hint): I prove the above claim $\mathbb{E}X = \mathbb{E}(-X)$. Notice that if U follows Unif(0,1) then $U' = 1-U$ also follows Unif(0,1) which is easily checked by the CDF of U'.
For $x \in (0,1)$, $\mathbb{P}(U' \le x) = \mathbb{P}(U \ge 1-x) = \int_{1-x}^1 1da = x$ (similarly you can check $x \le 0, \mathbb{P}(U' \le x) = 0; x \ge 1, \mathbb{P}(U' \le x) = 1;$); therefore, the CDF of U' is the distribution of Unif(0,1).
By LOTUS we know

$$\mathbb{E}(X) = \int_0^1 [\log(u/(1-u))] du \tag{2}$$

but at the same time we can rewrite X as $-\log((1-U)/U) = -\log(U'/(1-U'))$, and using the distribution of U' we would know that

$$\mathbb{E}(X) = \mathbb{E}(-\log(U'/(1-U'))) = -\int_0^1 [\log(u'/(1-u'))] du' = -\int_0^1 [\log(u/(1-u))] du \tag{3}$$

The left hand of (3) is equal to the minus the left hand of (2), and thus this implies that

$$\mathbb{E}(X) = \mathbb{E}(-X)$$

Therefore, we know

$$\mathbb{E}(X) = (\mathbb{E}(X) + \mathbb{E}(X))/2 = (\mathbb{E}(-X)\mathbb{E}(X))/2 = \mathbb{E}(0)/2 = 0$$

(2) Alternatively (second approach), we could derive the conclusion in an easier way (again use the fact if two r.v.s have the same distribution then they have the same expectation value): notice that if U and U' has the same distribution then by LOTUS we should have

$$\mathbb{E}\log(U) = \mathbb{E}\log(U')$$

as both sides are equal to $\int_0^1 \log(a)da$. Therefore:

$$\mathbb{E}X = \mathbb{E}\left(\log\frac{U}{U'}\right) = \mathbb{E}\left(\log U\right) - \mathbb{E}\left(\log U'\right) = 0$$

Further comment, Let's denote X, Y two r.v.s. with CDFs F_X, F_Y, notice that $X = Y$ implies $F_X = F_Y$, $F_X = F_Y$ implies $\mathbb{E}X = \mathbb{E}Y$ but the inverse may not hold. We can have $\mathbb{E}X = \mathbb{E}Y$ but $F_X \neq F_Y$ (think about Unif(-1,1), Unif(-2,2)) and we can have $F_X = F_Y$ but $X \neq Y$ (think about X and $1 - X$ in this exercise).

Normal

30. To show $S \sim N(0,1)$ we only need to show its distribution is the same as the one of Z (whose distribution is N(0,1)):

    $$\mathbb{P}\left(SZ \leq y\right) = \mathbb{P}\left(SZ \leq y|S=1\right)\mathbb{P}(S=1) + \mathbb{P}\left(SZ \leq y|S=1\right)\mathbb{P}(S=-1) = \mathbb{P}\left(Z \leq y\right)$$

    (or use MGF)

    Alternatively, we can use MGF to check:

    $$\begin{aligned} M_{SZ}(t) =& \mathbb{E}e^{tSZ} \\ =_{Independence}& \ \mathbb{E}e^{-tZ}1/2 + \mathbb{E}e^{tZ}1/2 = e^{t^2/2} \end{aligned}$$

    which is a MGF of N(0,1). (The above calculation of MGF is slightly beyond the scope of this course, as it may require the concept of conditional expectation.)

31. $\Phi(Z)$ $\sim$Unif(0,1), the rest is apparent.

This is the universality of the uniform, and read Theorem 5.3.1 in the textbook.

32.a. Symmetry of the normal implies:

$$\begin{aligned}\mathbb{P}\left(1 \leq X \leq 4\right) &= \mathbb{P}\left(1 \leq Z \leq 2 \ or \ -2 \leq Z \leq -1\right) \\ &= 2\mathbb{P}\left(1 \leq Z \leq 2\right)\end{aligned}$$

Note that $\mathbb{P}\left(1 \leq Z \leq 2\right) = \mathbb{P}\left(-\infty \leq Z \leq 2\right) - \mathbb{P}\left(-\infty \leq Z \leq 1\right) = \Phi(2) - \Phi(1)$. Now make use the following

$$\mathbb{P}\left(|X-\mu| < \sigma\right) \approx 0.68, \mathbb{P}\left(|X-\mu| < \sigma\right) \approx 0.95$$

and thus (by symmetry) $\Phi(2) = 0.5 + \mathbb{P}\left(|Z| < 2\right)/2 \approx 0.975$, $\Phi(1) = 0.5 + \mathbb{P}\left(|Z| < 1\right)/2 \approx 0.84$. Final result is $0.975 - 0.84 = 0.135$.

In general, notice

$$\Phi(t) = 0.5 + \mathbb{P}\left(|Z| < t\right)/2$$

32.b. Notice when $I_{Z>t} = 1$, $Z/t > 1$ and thus $I_{Z>t} \leq Z/tI_{Z>t}$ holds true.

$I_{Z>t} \leq Z/tI_{Z>t}$ implies that

$$\int_{-\infty}^{\infty} I_{z>t}\phi(z)dz \leq \int_{-\infty}^{\infty} z/tI_{z>t}\phi(z)dz \tag{4}$$

Note that $\int_{-\infty}^{\infty} I_{z>t}\phi(z)dz = 1-\Phi(t)$; $\int_{-\infty}^{\infty} z/tI_{z>t}\phi(z)dz = \int_{t}^{\infty} z/t\phi(z)dz = \int_{t}^{\infty} -1/td\phi(z) = \phi(z)/t|_{\infty}^{t}\phi(z) = 0) = \phi(t)/t$ (we use the fact ($\lim\limits_{z\to\infty}\phi(z) = 0$), and thus (4) implies

$$1 - \Phi(t) \leq \phi(t)/t$$

35.

$$\begin{aligned}\mathbb{E}\left(\max\{Z-c,0\}\right) &= \int_{\mathbb{R}} \max\{z-c,0\}\phi(z)dz \\ &= \int_{c}^{\infty} (z-c)\phi(z)dz \\ &= \int_{c}^{\infty} \phi(z)/2dz^2 - c(1-\Phi(c)) \\ &= \int_{c}^{\infty} -\frac{1}{\sqrt{2\pi}}de^{-z^2/2} - c(1-\Phi(c)) \\ &= \phi(c) - c(1-\Phi(c))\end{aligned}$$

If you fill in c=0, indeed we would have $1/\sqrt{2\pi}$ as suggested by the hint.

Now in case you are not familiar with the integration, take a look of the followings:

- $\int_c^\infty (z-c)\phi(z)dz = \int_c^\infty z\phi(z)dz - \int_c^\infty c\phi(z)dz$
- $\int_c^\infty c\phi(z)dz = \int_{-\infty}^\infty c\phi(z)dz - \int_{-\infty}^c c\phi(z)dz = c\int_{-\infty}^\infty \phi(z)dz - c\int_{-\infty}^c \phi(z)dz = c - c\Phi(c)$ *(by the fact that the integral of the PDF ϕ over the entire space is equal to one, and by definition of Φ, the CDF of standard normal distribution)*
- $\int_c^\infty z\phi(z)dz = \int_c^\infty z\frac{1}{2\sqrt{2\pi}}e^{-z^2/2}dz = \int_c^\infty -\frac{1}{\sqrt{2\pi}}(e^{-z^2/2})'dz = \left(-\frac{1}{\sqrt{2\pi}}(e^{-z^2/2})\right)\Big|_c^\infty - \int_c^\infty (-\frac{1}{\sqrt{2\pi}})'(e^{-z^2/2})$ *(integration by part).*

  Note that $(-\frac{1}{\sqrt{2\pi}})' = 0$ *and thus* $-\int_c^\infty(-\frac{1}{\sqrt{2\pi}})'(e^{-z^2/2}) = 0$; $\left(\frac{1}{\sqrt{2\pi}}(e^{-z^2/2})\right) = \phi(z)$, *and thus* $\left(\frac{1}{\sqrt{2\pi}}(e^{-z^2/2})\right)\Big|_c^\infty = -\ \phi(z)|_c^\infty = \lim_{x\to\infty}(-\phi(x)) - (-\phi(c)) = 0 - (-\phi(c)) = \phi(c)$.

Exponential

36.a. Hint: memoryless properties and symmetry. The result is 1/2.

Notice the probability that both Bob and Claire leave simultaneously has probability zero (as service time is continuous r.v.).

Without loss of generality, we consider the case Bob leaves first compared with Claire (this happens with probability 1/2 by symmetry, you can discuss the case Claire leaves first similarly), then conditional on that we know it is now a comparison between Claire and Alice. By memoryless property, we know the further service time needed for Claire and Alice will all follow identical exponential distribution, and thus by symmetry we know Alice will leave after Claire with probability 1/2.

36.b. **The expected waiting time before the service** $1/(2\lambda)$, since the minimum of two independent Expo(λ) follows Expo(2λ) (Check Example 5.6.3 for a proof of this statement which proves the statement by checking the CDF.).

Then the her own service's expected duration is $1/\lambda$ as it follows Expo(λ).

Result: $1/\lambda + 1/(2\lambda)$.

43.a. Assume the first trial happens at time position 0, then the jth trial would occur at $(j-1)\Delta t$. Thus

$$T = ((G+1)-1)\Delta t$$

since the first success trial would be the $(G+1)$th trial.

43.b. For $t \geq 0$

$$\mathbb{P}(T > t) = \mathbb{P}(G > \lfloor t/(\Delta t) \rfloor) = (1-\lambda\Delta t)^{\lfloor t/(\Delta t)\rfloor + 1}$$

where $\lfloor x \rfloor = \max\{n \in \mathbb{N} : n \leq x\}$. Thus for $t \geq 0$

$$\mathbb{P}(T \leq t) = 1 - (1-\lambda\Delta t)^{\lfloor t/(\Delta t)\rfloor + 1}$$

For $t < 0, \mathbb{P}(T \leq t) = 0$**.**

43.c. From $(t/(\Delta t) - 1) \geq \lfloor t/(\Delta t) \rfloor \geq t/(\Delta t)$ we know

$$(1-\lambda\Delta t)^{t/(\Delta t)} / (1-\lambda\Delta t) \geq (1-\lambda\Delta t)^{\lfloor t/(\Delta t)\rfloor+1} \geq (1-\lambda\Delta t)^{t/(\Delta t)+1} \tag{5}$$

Note that $\lim_{\Delta t \to 0} (1-\lambda\Delta t) = 1$ and

$$\lim_{\Delta t \to 0} \left[(1-\lambda\Delta t)^{-1/(\lambda\Delta t)}\right]^{-\lambda t} = e^{-\lambda t}$$

Therefore, taking limits on the both sides of equation (5) we know $(1-\lambda\Delta t)^{\lfloor t/(\Delta t)\rfloor+1}$ converges to $e^{-\lambda t}$ and thus $\mathbb{P}(T \leq t)$ converges to $1 - e^{-\lambda t}$ (the Expo(λ) CDF).

45. Notice $T > 0.1$ means at most 2 arrives $N \leq 2$:

$$\mathbb{P}(T > 0.1) = \mathbb{P}(N \leq 2) = e^{-2}\left(2^0/0! + 2^1/1! + 2^2/2!\right) = 5e^{-2}$$

(The arrival rate within 0.1 hour is 2).

Mixed practice

51.a. Use the hint we know $\mathbb{E}X^2 \leq \mathbb{E}X$, and thus

$$\mathbb{V}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 \leq \mathbb{E}X - (\mathbb{E}X)^2 = \mu(1-\mu) = -(\mu - 1/2)^2 + 1/4 \leq 1/4$$

First let's prove that the hint. we have seen one example already to prove one random inequality holds true with probability one in the tutorial (6.12.a). Here we prove $X^2 \leq X$ similarly.

To derive $\mathbb{P}(X^2 \leq X) = 1$, we essentially are proving that for all possible numbers taken by X, for all possible outcomes of X, (in other words, for all elements in the sample space as X maps these elements into numbers) the inequality holds. Namely, for an arbitrary number x in the range of X we need to have $x^2 \leq x$. X only takes values in the interval $(0, 1)$, and for $x \in (0, 1)$ we know $x^2 \leq x$. Therefore, we know the even $\{X^2 \leq X\}$ happens with probability equal to one.

Then given $\mathbb{P}(X^2 \leq X) = \mathbb{P}(X - X^2 \geq 0) = 1$, we know for any arbitrary large number $M > 2$ (I use M to control the bound of $X - X^2$ during the even $\{X - X^2 \leq 0\}$, (note that $|X - X^2| \leq |X| + |X^2| \leq 1$ as $0 \leq X \leq 1$), only for mathematical rigor, and normally you can also simply ignore what happens in these zero-probability events (would be zero anyway).)

$$\mathbb{E}(X - X^2) \geq 0 \times \mathbb{P}(X - X^2 \geq 0) - M \times \mathbb{P}(X - X^2 \leq 0) = 0 - M \times 0 = 0$$

Therefore, $\mathbb{E}X^2 \leq \mathbb{E}X$ and thus

$$\mathbb{V}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 \leq \mathbb{E}X - (\mathbb{E}X)^2 = \mu(1 - \mu) \leq -(\mu - 1/2)^2 + 1/4 \leq 1/4$$

Further comment: we should be so surprised to see random inequalities which may hold with certain probabilities (sometimes one, sometimes not).

We have already seen other cases such that for X, Y independent identical distributed continuous r.v.s we have $\mathbb{P}(X > Y) = 1/2$ by symmetry.
We discuss $\mathbb{P}(X > Y) = 1/2$ here again to enhance your understanding. We can partition sample space (or events) into three cases: $\{X > Y\}$ denotes elements in the sample space that result in $X > Y$; $\{Y < X\}$ and $\{X = Y\}$. We know $\mathbb{P}(X = Y) = 0$ since X, Y are continuous r.v.s, while by symmetry we know $\{X > Y\}$ and $\{Y < X\}$ should occur with equal probability.

51.b. Bern(1/2). Hint: The equality of ($\mathbb{E}X^2 - (\mathbb{E}X)^2 \leq \mathbb{E}X - (\mathbb{E}X)^2$) holds only if $X^2 = X$.

Based on the above discussion in (a), we know $\mathbb{V}X = 1/4$ needs $\mathbb{E}(X) = \mathbb{E}(X^2)$ (otherwise it is strictly smaller that 1/4) and when $0 \leq X \leq 1$, $X^2 \leq X$ holds true with probability equal to one.

We decompose the $\{X^2 \leq X\}$ into two disjoint events: $\{X^2 < X\}$ and $\{X^2 = X\}$. If $\mathbb{P}(X^2 < X) = c > 0$, then we know $\mathbb{E}(X) > \mathbb{E}(X^2)$

$$\mathbb{E}(X - X^2) > 0 \times \mathbb{P}(X - X^2 \geq 0) - M \times \mathbb{P}(X - X^2 \leq 0) = 0 \times c - M \times 0 = 0$$

Therefore, to have $\mathbb{E}(X) = \mathbb{E}(X^2)$ we must have $\mathbb{P}(X^2 < X) = 0$ and $\mathbb{P}(X^2 = X) = 1 - \mathbb{P}(X^2 < X) = 1$.

The only non-degenerated r.v. X such that $X^2 = X$ (so X can only take 1 or 0) with probability equal to one would be Bern(p). For Bern(p) we know the variance is $p(1-p)$ and we know p can only be 1/2 in order to have $p(1-p) = 1/4$.

55.a. Hint: $J - 1 \sim$Geom$(1 - \Phi(2))$. Thus $1/(1 - \Phi(2))$.

(Note that in this exercise, you are drawing X_i's from a normal distribution, and each X_i can be any values in $\mathbb{R}$, and you only succeed if you happen to draw one X_i that is strictly larger than 4. For each draw you have a successful rate equal the probability of X_i being strictly larger than 4.)

Note that the successful rate for each draw is $\mathbb{P}\left(X_i > 4\right) = \mathbb{P}\left((X_i - 0)/\sqrt{4} > (4-0)/\sqrt{4}\right) = 1 - \Phi(2)$, as $(X_i - 0)/\sqrt{4} \sim N(0,1)$. The we know $J - 1 \sim$Geom$(1 - \Phi(2))$ and thus $\mathbb{E}J = 1/(1 - \Phi(2))$ (we use the result of Example 4.3.7 to derive the expectation).

55.b. 1. (Simply integration)

By LOTUS,

$$\mathbb{E}\frac{g(X)}{f(X)} = \int_{-\infty}^{\infty} \frac{g(x)}{f(x)} f(x) ds = \int_{-\infty}^{\infty} g(x) ds = 1$$

The last equality is true because g is one PDF (which needs to have integration over all values equal to 1 by definition).

Further comment (beyond the scope of the course): now if we take a closer look at the R ratio, which is a ratio between f and g, by using this R, we can approximate $\mathbb{E}g(X)$ with X's PDF being f, by independent N random draws y_i from one arbitrary distribution g as long as the support of g is larger than f we can have $\sum_i H(y_i)R(y_i)/N$. This is quite useful in practice, as computer can simulate random draws from certain distributions (unfortunately not all distributions, and that is why we have the ratio R here) quite easily.

55.c. $W \sim$Unif(0,1). (The rests are easy.)

By universality of the uniform (theorem 5.3.1) we know $W \sim$Unif(0,1) since $F(x)$ is strictly increasing in x with range equal to [0,1]. Thus we know mean is 1/2 and variance is 1/12 by properties of Unif(0,1) (we have also calculated these during lectures, you can also check example 5.2.4).

The alternative way is to derive the CDF of $F(X)$: for $x \in [0,1]$

$$\mathbb{P}(F(X) \leq x) = \mathbb{P}(X \leq F^{-1}(x)) = F\left(F^{-1}(x)\right) = x$$

where the last equality is due to the fact that $X's$ CDF function is F. Therefore, from $x = 0$ to 1 $\mathbb{P}(F(X) \leq x)$ increases linearly from 0 to 1, and this is the CDF of Unif(0,1). Thus we know mean is 1/2 and variance is 1/12 by properties of Unif(0,1) (we have also calculated these during lectures, you can also check example 5.2.4).

56.a. $\int_{\mathbb{R}} z^2\Phi(z)\phi(z)dz$

Note here we need to calculate the expectation of one r.v. which is a transformation of r.v. Z (Φ is just one function!! Note that location transformation is a specific linear transformation of one r.v., you could consider the location-scale transformation as a transformation of a given r.v. X with mean μ variance σ^2 using the function $f(a) = (a-\mu)/\sigma$). By LOTUS, we only need to know the distribution of Z to calculate the expectation (no need to know the distribution of the transformation $Z^2\Phi(Z)$):

$$\mathbb{E}Z^2\Phi(Z) = \int_{\mathbb{R}} z^2\Phi(z) \underbrace{\phi(z)}_{\text{The PDF of Z}} dz$$

56.b. 2/3.

By Universality of the Uniform, $\Phi(Z) \sim$Unif(0,1):

$$\mathbb{P}\left(\Phi(Z) < 2/3\right) = 2/3$$

Alternatively, just calculate the probability directly:

$$\mathbb{P}\left(\Phi(Z) < 2/3\right) = \mathbb{P}\left(Z < \Phi^{-1}(2/3)\right) = \Phi\left(\Phi^{-1}(2/3)\right) = 2/3$$

which uses the steps in the proof of Universality of the Uniform.

Remark: Universality of the Uniform simply tells us the CDFs of certain transformed r.v. For many transformations, we do not know the distributions of the transformations, but for some we know. For example, we know standardized transformation of one arbitrary normal distributed r.v. will have standard normal distribution, and Universality of the Uniform provides two more cases.

56.c. symmetry of continuous r.v.s. (Result: 1/3!).

This is essentially proposition 5.7.1. Let's order three r.v.s, there would be 3! different orderings (e.g., one of them could be $Y < Z < X$), and by symmetry, all orderings of X, Y, Z are with equal probability. Further more we know the probability of any of these two r.v.s are equal would be zero, as they are all independent **continuous** r.v.s.

Therefore, we know all orderings of X, Y, Z are with equal probability and their sum is equal to one:

$$\mathbb{P}(X < Y < Z) = 1/3!$$