Solution to Example 5

(a) It is obvious that $f(x) > 0$ for $x > 0$.

$$\int_{-\infty}^{\infty} f(x;\theta)\,dx = \int_0^{\infty} \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}}\,dx$$
$$= \int_0^{\infty} d\left(e^{-\frac{x^2}{2\theta^2}}\right)$$
$$= \left[e^{-\frac{x^2}{2\theta^2}}\right]_0^{\infty}$$
$$= 0 - (-1) = 1.$$

A small amendment:

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x;\theta)dx &= \int_0^{\infty} \frac{x^2}{\theta} e^{-\frac{x^2}{2\theta^2}} dx \\
&= -\int_0^{\infty} d\left(e^{-\frac{x^2}{2\theta^2}}\right) \\
&= -\left[e^{-\frac{x^2}{2\theta^2}}\right]_0^{\infty} \\
&= -[0-1] = 1.
\end{aligned}$$

Remarks

1. The expected value exists provided the sum or the integral in the above definitions exists.

2. In the discrete case, if $f_X(x) = 1/N$ for each of the N values of x, hence the mean,

$$E(X) = \sum_i x_i\, f(x_i) = \frac{1}{N}\sum_i x_i\,,$$

becomes the average of the N items.

We give one example that the expectation does not exist:

Assume that X takes point masses on $1^2, 2^2, 3^2, \ldots,$; and for each $n = 1, 2, \ldots,$

$$P(X = n^2) = \frac{6}{\pi^2}\frac{1}{n^2}.$$

Note that it is well known $\sum_{n=1}^{\infty}\frac{1}{n^2} = \pi^2/6$, which is called the Basal's problem, and the solution was first found by Euler in 1735; therefore, this gives an appropriately defined (discrete) probability distribution. However, if we try to evaluate $E(X)$ using the formula given in page 2-87, we have

$$E(X) = \sum_{n=1}^{\infty} n^2 \frac{1}{n^2}\frac{6}{\pi^2} = \frac{6}{\pi^2}\sum_{n=1}^{\infty} 1 = \infty.$$

Please try on your own to find an example that the expectation of a continuous random variable does not exist.

Pages 2-91 & 92

Example 1

- In a gambling game, a man gains 5 if he gets all heads or all tails in tossing a fair coin 3 times,
- and he pays out 3 if either 1 or 2 heads show.
- What is his expected gain?

Solution to Example 1

- Let X be the amount he can gain.
- Then $X = 5$ or -3 with the following probabilities:

  $\Pr(X = 5) = \Pr(\{HHH, TTT\}) = 1/8 + 1/8 = 1/4$

  $\Pr(X = -3) = 1 - \Pr(X = 5) = 3/4.$
- Therefore $E(X) = 5\left(\frac{1}{4}\right) + (-3)\left(\frac{3}{4}\right) = -1.$
- Hence, he will lose 1 per toss **in a long run**.

How could we change the pay amount in 2nd item “and he pays 3 if either 1 or 2 heads show” on page 2-91 so that the game is a fair game?

Here fair means that the expected gain would be equal to 0.

We can set the amount he pays out if either 1 or 2 heads show is a. Then the expected gain would be

$$E(X) = 5(1/4) + (-a)(3/4)$$

Then setting this expected value to be 0 and solve the equation for a, we can get $a = 5/3$, which is the amount to replace 3 so that the game is fair.

Certainly 5/3 is not a practical amount to pay in a single game. Is there any other way that we can adjust to make the game fair? Try to adjust the gains when all heads and all tails are obtained in the game!

Pages 2-93 & 94

Example 2

- Suppose a game consists of rolling a balanced die.
- We pay c to play the game and we get i if number i occurs.
- How much should we pay if the game is fair?
  (We say a game is "fair" if $E(\text{gain}) = 0$.)

Solution to Example 2

- Let X denote the amount that one gets when rolling a die.
- Then clearly $\Pr(X = 1) = \cdots = \Pr(X = 6) = 1/6$.
- $E(X) = (1 + 2 + \cdots + 6)\left(\frac{1}{6}\right) = 3.5$.
- To be a fair game, E(paying) = E(getting), and hence for a fair game, the admission fee should be $c = E(X) = 3.5$.

Please get familiar with the computation of the expectation for discrete random variables; it is the "weighted average" of the possible values of the random variable, where "weights" are the corresponding probabilities.

In this example, if the die is unfair, such that the probability of getting 1 is 0.5, and getting 2, 3,4,5,6 are all 0.1, then

$$E(X) = 0.5 \cdot 1 + 0.1 \cdot (2 + 3 + 4 + 5 + 6) = 2.5.$$

So to ensure that the game is fair, we need to set $c = 2.5$, which is lower than 3.5. This intuitively makes sense. As we have greater chance to get the lowest reward 1, to balance such a results, we need to pay less to play the game so that we won't "expect" to lose in the long run.

2.5.2 Expectation of a Function of a RV

Definition 2.8

For any function $g(X)$ of a random variable X with p.f. (or p.d.f.) $f_X(x)$,

(a) $E[g(X)] = \sum_x g(x) f_X(x)$

if X is a **discrete** r.v. providing the sum exists; and

(b) $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

if X is a **continuous** r.v. providing the integral exists.

This page gives very important and useful formulae. These formulae allow us to compute the expectations of an arbitrary function of a random variable with a known distribution.

For example, if we have a continuous random variable X, we know its p.d.f. is $f_X(x)$, and we want to find $E(\sin(X))$, we do not need to derive the p.d.f. for $\sin(X)$, instead, we only need to apply the formula $E(\sin(X)) = \int_{-\infty}^{\infty} \sin(X) f_X(x) dx$ to compute this expectation.

Some discussion on the expectation $E(X)$ and variance $V(X)$:

✓ Expectation is essentially given the "population mean" (or intuitively the "central location" of the possible values) of X. Therefore, one may also see that $E(X)$ is called the location parameter in some literature.

✓ Variance tells how the possible values of X spread around $E(X)$. The smaller the $V(X)$, the more concentrated the possible values of X around $E(X)$; and vice versa. $V(X)$ is frequently called the dispersion parameter in many places.

✓ Note that $V(X)$ has two alternative (but in fact equivalent) formulae to get:

$$V(X) = E\left((X - E(X))^2\right) = E(X^2) - (E(X))^2 . \tag{1}$$

★ The former serves as the definition of variance given on page 2-105.

Some Special Cases (Continued)

Definition 2.9

- Let X be a random variable with p.f. (or p.d.f.) $f(x)$, then the **variance** of X is defined as

$$\sigma_X^2 = V(X) = E[(X - \mu_X)^2]$$
$$= \begin{cases} \displaystyle\sum_x (x - \mu_X)^2 f_X(x), & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx, & \text{if } X \text{ is continuous.} \end{cases}$$

ST2334 Probability and Statistics CYM Concepts of Random Variables 2-105

★ The latter is mathematically derived on page 2-126.

Properties of Expectation (Continued)

Property 2 (Continued)

Proof:

$$V(X) = E[(X-\mu_X)^2]$$
$$= E[X^2 - 2X\mu_X + \mu_X^2]$$
$$= E(X^2) - E(2X\mu_X) + E(\mu_X^2)$$
$$= E(X^2) - 2\mu_X\, E(X) + (\mu_X^2)$$
$$= E(X^2) - 2\mu_X^2 + \mu_X^2 = E(X^2) - \mu_X^2$$

Notice that $\mu_X = E(X)$ is a constant.

★ The latter formula is usually computationally more convenient when we have the p.d.f. available and we are to compute the variance. Examples 1–3 on page 2-108 to 2-121 give good illustrate of this.

✓ One very obvious observation based on (1) is that $E(X^2) \geq (E(X))^2$, and "=" can hold if and only if X is a constant (i.e., nonrandom); more specifically, there is a given value c, such that $P(X = c) = 1$. In such a case, $V(X) = 0$ (i.e., no variability for X).

✓ Several very useful formulae for expectation and variance are given below:

★ For the expectation, it has the linearity: for ANY random variables $X_1, X_2, \ldots, X_i$, and constants (nonrandom) $a_1, a_2, \ldots, a_k$,

$$E(a_1X_1 + a_2X_2 + \ldots + a_kX_k) = a_1E(X_1) + a_2E(X_2) + \ldots + a_kE(X_k).$$

This is a slightly more general formula than that given on page 2-124. Think about why.

Properties of Expectation (Continued)

Two special cases:

(a) Put $b = 0$, we have $E(aX) = a\,E(X)$.

(b) Put $a = 1$, we have $E(X + b) = E(X) + b$.

In general,

$$E[a_1 g_1(X) + a_2 g_2(X) + \cdots + a_k g_k(X)]$$
$$= a_1 E[g_1(X)] + a_2 E[g_2(X)] + \cdots + a_k E[g_k(X)]$$

where $a_1, a_2, \cdots, a_k$ are constants.

★ For the variance, we have the formula:

$$V(aX + b) = V(aX) = a^2 V(X),$$

which can be viewed as the combination of two formulae:

- for any random variable Y and constant b, $V(Y + b) = V(Y)$;

$$V(Y + b) = E\left(((Y + b) - E(Y + b))\right) = E\left((Y - E(Y))^2\right) = V(Y);$$

- and for any constant a and random variable X, $V(aX) = a^2 V(X)$;

$$V(aX) = E\left((aX - E(aX))^2\right) = a^2 E\left((X - E(X))^2\right) = a^2 V(X).$$

The development above can also be used to replace the proof given on page 2-128 of the lecture slides.

Properties of Expectation (Continued)

Property 3 (Continued)

Proof:

$$
\begin{aligned}
V(aX+b) &= E[(aX+b)^2] - [E(aX+b)]^2 \\
&= E(a^2X^2 + 2abX + b^2) - (a\mu_X + b)^2 \\
&= a^2E(X^2) + 2ab\,E(X) + b^2 - (a^2\mu_X^2 + 2ab\mu_X + b^2) \\
&= a^2E(X^2) - a^2\mu_X^2 \\
&= a^2[E(X^2) - \mu_X^2] \\
&= a^2V(X).
\end{aligned}
$$

ST2334 Probability and Statistics CYM Concepts of Random Variables 2-128

✓ In practice, we may jointly consider the expectation and the variance to judge the performance of a random variable. See the example given on page 2-129. The expected profit (700$) derived on page 130 looks promising. The standard deviation (800$ given on page 2-131), however, is much too big. Eventually, the chance that the shop will lose money is not small (for this specific example, you can make scenarios that the shop may lose money; you can also compute the probability that the shop will lose money on your own), especially in a short period of time. So, if you are the boss of the shop, you might think of whether you want to take the risk.

Example 4

- A jewelry shop purchased three necklaces of a certain type at \$500 a piece.
- It will sell them for \$1000 a piece. The designer has agreed to repurchase any necklace still unsold after a specified period at \$200 a piece.
- Let X denote the number of necklaces sold and suppose X follows the following probability distribution.

x	0	1	2	3
$f_X(x)$	0.1	0.2	0.3	0.4

- Find the expected gain and the variance of the gain.

Solution to Example 4

- With $g(X) = \text{revenue} - \text{cost} = 1000X + 200(3 - X) - 3(500) = 800X - 900$.
- $E\big(g(X)\big) = g(0)f_X(0) + g(1)f_X(1) + g(2)f_X(2) + g(3)f_X(3)$
  $= (-900)(0.1) + (-100)(0.2) + (700)(0.3) + 1500(0.4)$
  $= 700.$
- Hence the expected profit is \$700.

Solution to Example 4 (Continued)

- $E\left[\left(g(X)\right)^2\right] = (-900)^2(0.1) + (-100)^2(0.2) + (700)^2(0.3) + 1500^2(0.4) = 1130000$.

- Hence
$$V\left(g(X)\right) = 1130000 - 700^2 = 640000$$
and
$$\sqrt{V\left(g(X)\right)} = \sqrt{640000} = 800$$

Chebyshev's Inequality (Continued)

- Let X be a random variable (discrete or continuous) with $E(X) = \mu$ and $V(X) = \sigma^2$.
- Then for any positive number k we have

$$\Pr(|X - \mu| \geq k\sigma) \leq 1/k^2.$$

- That is, the probability that the value of X lies at least k standard deviation from its mean is at most $1/k^2$.
- Alternatively,

$$\Pr(|X - \mu| < k\sigma) \geq 1 - 1/k^2.$$

✓ Make clear that Chebyshev's inequality only provides some bounds for the probabilities that the random variable will take a values in a certain range; these are not sharp bounds in most cases.

On the other hand, it definitely plays fundamental roles in the developments of probability and statistical theories, though these are beyond the scope of this module.

✓ A slightly more handy forms of these formulae are given below

$$\begin{aligned} Pr(|X - \mu| \geq c) &\leq \frac{V(X)}{c^2} \\ Pr(|X - \mu| < c) &\geq 1 - \frac{V(X)}{c^2}, \end{aligned}$$

for any constant $c > 0$. Bear in mind these are exactly the same formulae as those given in the slide.