4. Further Topics on Random Variables

Han-You Jeong

School of Electrical and Computer Engineering

Pusan National University

Probability and Statistics

Outline

Chapter Overview

- Introduction to methods that are useful in:
  1. deriving the distribution of a function of random variable(s);
  2. dealing with the sum of independent random variables;
  3. quantifying the degree of dependence between two random variables.

- Introduction to a number of tools including transforms and convolutions.

Calculation of the PDF of a Function $Y = g(X)$

- We wish to calculate the PDF of a function $Y = g(X)$ of a continuous R. V. X
  1. Calculate the CDF F_Y of Y using the formula
  $$F_Y(y) = P(g(X) \leq y) = \int_{\{x|g(x)\leq y\}} f_X(x)dx.$$
  2. Differentiate to obtain the PDF of Y
  $$f_Y(y) = \frac{dF_Y(y)}{dy}.$$

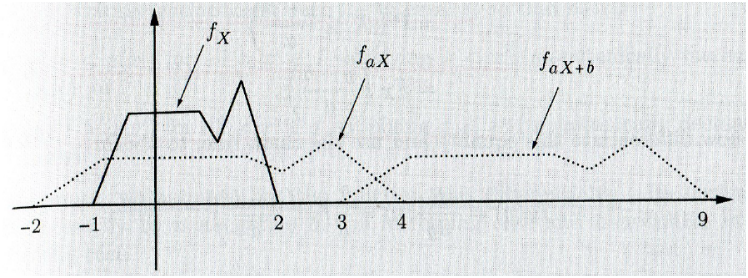
Ex 4.2 John is driving from Boston to the New York area, a distance of 180 Km at a constant speed, whose value is uniformly distributed between 30 and 60 Km/h. What is the PDF of the duration of the trip?

Ex 4.3 Let $Y = g(X) = X^2$, where X is a R. V. with known PDF $f_X(x)$. What is $f_Y(y)$ in terms of f_X?

The Linear Case $Y = aX + b$

- Let X be a continuous R. V. with PDF f_X, and let $Y = aX + b$, where a and b are scalars with $a \neq 0$. Then,

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$



Ex 4.4 Suppose X is an exponential R. V. with parameter λ. Let $Y = aX + b$, then find the PDF of Y.

The Monotonic Case

- Let X be a continuous R. V. and suppose that its range is contained in an interval I, in the sense that $f_X(x) = 0$ for $x \notin I$. For the R. V. $Y = g(X)$, function g is called strictly monotonic over the interval I, either
  1. $g(x) < g(x')$ for all $x, x' \in I$ s.t. $x < x'$ (monotonically increasing); or
  2. $g(x) > g(x')$ for all $x, x' \in I$ s.t. $x < x'$ (monotonically decreasing).
- Suppose that g is strictly monotonic and that for some function h and all x in the range of X we have

$$y = g(x) \quad \text{if and only if} \quad x = h(y).$$

Assume that h is differentiable, the PDF of Y in the region where $f_Y(y) > 0$ is given by

$$f_Y(y) = f_X(h(y)) \left| \frac{dh(y)}{dy} \right|.$$

Ex 4.6 Let $Y = g(X) = X^2$, where X is a continuous uniform R. V. on the interval [0, 1]. Find the PDF of R. V. Y.

Functions of Two Random Variables

- The two-step procedure that first calculate the CDF and then differentiates to obtain PDF also applies to functions of more than one R. V.

  Ex 4.7 Two archers shoot at a target. The distance of each shot from the center is uniformly distributed from 0 to 1, independent of the other shot. What is the PDF of the distance of the losing shot from the center?

  Ex 4.9 Romeo and Juliet have a date at a given time, and each, independently, will be late by an amount of time that is exponentially distributed with parameter λ. What is the PDF of the difference between their times of arrival?

Sum of Independent Random Variables - Convolution

- Let $Z = X + Y$, where X and Y are independent discrete R. V.s with PMFs p_X and p_Y, respectively. Then, for any z, we have

$$p_Z(z) = P(X+Y=z) = \sum_{\{(x,y)|x+y=z\}} P(X=x, Y=y)$$
$$= \sum_{\forall x} P(X=x, Y=z-x) = \sum_{\forall x} p_X(x)p_Y(z-x),$$

  which is called the convolution of the PMFs of X and Y.
- Suppose that X and Y are independent continuous R. V.s with PDFs f_X and f_Y,respectively. Then, we have

$$P(Z \leq z|X=x) = P(X+Y \leq z|X=x) = P(x+Y \leq z|X=x)$$
$$= P(x+Y \leq z) = P(Y \leq z-x).$$

- By differentiating both sides, we have

$$f_{Z|X}(z|x) = f_Y(z-x).$$

- Using multiplication rule, we have

$$f_{X,Z}(x,z) = f_X(x) f_{Z|X}(z|x) = f_X(x) f_Y(z-x).$$

Finally, we obtain

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Z}(x,z)dx = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx.$$

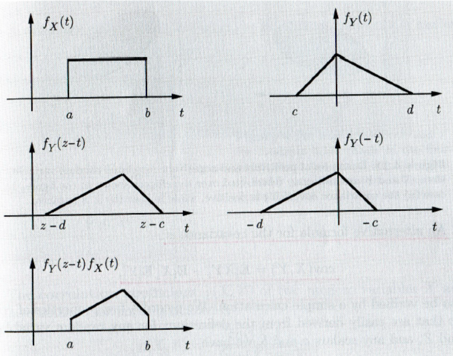
Ex 4.10 Two R. V.s X and Y are independent and uniformly distributed in the interval [0, 1]. What is the PDF of $Z = X + Y$?

Ex 4.12 Two R. V.s X and Y are independent and exponentially distributed with parameter λ. What is the PDF of $Z = X - Y$?

Graphical Calculation of Convolutions

- Consider two PDFs $f_X(t)$ and $f_Y(t)$. For a fixed value of z, the graphical evaluation of the convolution is as follows:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t) f_Y(z-t)\, dt$$



Correlation and Covariance

- The correlation of two R. V.s X and Y is defined by

$$cor(X,Y) = E[XY] = \begin{cases} \sum_{\forall x}\sum_{\forall y} xyp_{X,Y}(x,y), & \text{discrete R. V.;} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy, & \text{continuous R. V.} \end{cases}$$

  - If $cor(X,Y) = E[X]E[Y]$, two R. V.s X and Y are uncorrelated.
  - If $cor(X,Y) = 0$, two R. V.s X and Y are orthogonal.
- The covariance of two R. V.s X and Y is defined by

$$cov(X,Y) = E\left[(X - E[X])(Y - E[Y])\right] = E[XY] - E[X]E[Y].$$

  - For any R. V.s X, Y, and Z, and any scalars a and b, we have

$$\begin{aligned} &cov(X,X) = var(X), \\ &cov(X,aY+b) = a \cdot cov(X,Y), \\ &cov(X,Y+Z) = cov(X,Y) + cov(X,Z). \end{aligned}$$

Correlation Coefficient

Ex 4.13 The pair of R. V.s (X, Y) takes the values (1, 0), (0, 1), (-1, 0), and (0, -1), each with probability 1/4. Are these R. V.s uncorrelated and/or independent?

- The correlation coefficient $\rho(X, Y)$ of two R. V.s X and Y that have nonzero variances is defined as

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}}.$$

The correlation coefficient is normalized in the sense that $\rho(X, Y)$ ranges from -1 to 1.

Ex 4.14 Consider n independent tosses of a coin with probability of a head equal to p. Let X and Y be the number of heads and tails, respectively. Find the correlation coefficient $\rho(X, Y)$.

Variance of the Sum of Random Variables

- The covariance is used to obtain a formula for the variance of the sum of several (not necessarily independent) R. V.s.
- If $X_1, X_2, \cdots, X_n$ are R. V.s with finite variance, we have

$$var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} var(X_i) + \sum_{\{(i,j)|i\neq j\}} cov(X_i, X_j).$$

Ex 4.15 n people throw their hats in a box and then pick a hat at random. Find the variance of X, the number of people who pick their own hat.

Law of Iterated Expectations

- Since $E[X|Y=y]$ is a function of y, $E[X|Y]$ is also a function of Y and its distribution is determined by the distribution of Y.
  - We view the conditional expectation $E[X|Y]$ as a random variable determined by Y
  - The expectation of this random variable $E[E[X|Y]]$ is formulated by

$$E[E[X|Y]] = \begin{cases} \sum_y E[X|Y=y]p_Y(y), & Y \text{ discrete} \\ \int_{-\infty}^{\infty} E[X|Y=y]f_Y(y)dy, & Y \text{ continuous.} \end{cases}$$

- From the total expectation theorem, the law of iterated expectations $E[E[X|Y]] = E[X]$ holds.

Ex 4.17 We break a stick of length l at a point chosen randomly and uniformly, and keep the piece that contains the left end of the stick. We then repeat this process once again. What is the expected length of the piece that we are left with?

The Conditional Expectation as an Estimator

- If we view Y as an observation that provides information about X, it is natural to view the conditional expectation $\hat{X} = E[X|Y]$ as an estimator of X given Y.
- The estimation error $\tilde{X} = \hat{X} - X$ is a random variable satisfying

$$E[\tilde{X}|Y] = E[\hat{X} - X|Y] = E[\hat{X}|Y] - E[X|Y] = \hat{X} - \hat{X} = 0$$

  - The random variable $E[\tilde{X}|Y]$ is identically zero, and $E[\tilde{X}] = E[E[\tilde{X}|Y]] = 0$ by the law of iterated expectations. In addition,

$$E[\hat{X}\tilde{X}] = E[E[\hat{X}\tilde{X}|Y]] = E[E[\hat{X}(\hat{X} - X)|Y]] = E[\hat{X}^2 - \hat{X}E[X|Y]] = 0,$$

- The estimator $\hat{X}$ is uncorrelated with the estimation error $\tilde{X}$

$$cov(\tilde{X}, \hat{X}) = E[\hat{X}\tilde{X}] - E[\hat{X}]E[\tilde{X}] = 0$$

- Since $X = \hat{X} + \tilde{X}$, we have $var(X) = var(\hat{X}) + var(\tilde{X})$.

Law of Total Variance

- We introduce the random variable

$$var(X|Y) = E[(X - E[X|Y])^2|Y] = E[\tilde{X}^2|Y].$$

- Using the fact $E[\tilde{X}] = 0$ and the law of iterated expectations, we have

$$var(\tilde{X}) = E[\tilde{X}^2] = E[E[\tilde{X}^2|Y]] = E[var(X|Y)].$$

- Finally, the law of total variance states that

$$var(X) = var(\tilde{X}) + var(\hat{X}) = E[var(X|Y)] + var(E[X|Y]).$$

Ex 4.17 (continued) What is the variance of the length of piece that we are left with?

Homework # 4

- Section 4.1 Derived Distributions: Examples 4.9 and 4.12, Problems 7, 11, and 14
- Section 4.2 Covariance and Correlation: Problem 18
- Section 4.3 Conditional Expectation and Variance Revisited: Examples 4.18 and 4.20, Problems 22 and 24