Probability Theory

Chapter 7: Properties of Expectation

Lecturer

Dr. Monjed H. Samuh

Applied Mathematics & Physics Department
Palestine Polytechnic University
(monjedsamuh@ppu.edu)

Term 191

Table of Contents

Introduction

- The expected value of the random variable X is defined by

$$E(X) = \begin{cases} \sum_x xp(x) & \text{if } X \text{ is a discrete random variable} \\ \int\limits_{-\infty}^{\infty} xf(x)dx & \text{if } X \text{ is a continuous random variable} \end{cases},$$

Theorem

If $P(a \leq X \leq b) = 1$, then $a \leq E(X) \leq b$.

Expectation of Sums of Random Variables

Suppose that X and Y are random variables and g is a function of two random variables.

- If X and Y have a joint probability mass function $p(x, y)$, then

$$E(g(X,Y)) = \sum_y \sum_x g(x,y)p(x,y).$$

- If X and Y have a joint probability density function $f(x, y)$, then

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy.$$

- Suppose that $E(X)$ and $E(Y)$ are both finite and let $g(X, Y) = X + Y$. Then, in the continuous case,

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)f(x,y)dxdy = E(X)+E(Y).$$

Expectation of Sums of Random Variables

Example

Suppose that for random variables X and Y, $X \geq Y$. Then $E(X) \geq E(Y)$.

- $X \geq Y$ means for any outcome of the probability experiment, the value of the random variable X is greater than or equal to the value of the random variable Y.
- $X \geq Y$ is equivalent to $X - Y \geq 0$.
- It follows that $E(X - Y) \geq 0$.
- or, equivalently, $E(X) \geq E(Y)$.

- We may show by simple induction proof that if $E(X_i)$ is finite for all $i = 1, 2, \ldots, n$, then

$$E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n).$$

Expectation of Sums of Random Variables

Example (The sample mean)

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having distribution function F and expected value μ. Such a sequence of random variables is said to constitute a sample from the distribution F. The quantity

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

is called the sample mean. Compute $E(\bar{X})$.

Example (Boole's inequality)

- Let $A_1, A_2, \ldots, A_n$ denote events, and define the indicator variables $X_i,\ i = 1, 2, \ldots, n$ by

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases},$$

- Let $X = \sum_{i=1}^{n} X_i$ denotes the number of the events A_i that occur.

- Let

$$Y = \begin{cases} 1 & \text{if } X \geq 1 \\ 0 & \text{otherwise} \end{cases},$$

  so Y is equal to 1 if at least one of the A_i occurs and is 0 otherwise.

Expectation of Sums of Random Variables

Example (Boole's inequality - Cont.)

- Now, it is immediate that $X \geq Y$.
- $E(X) \geq E(Y)$.
- $E(X) = \sum_{i=1}^{n} P(A_i)$.
- $E(Y) = P\left(\bigcup_{i=1}^{n} A_i\right)$.
- We obtain Boole's inequality, namely,

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

Expectation of Sums of Random Variables

Example (Expectation of a binomial random variable)

- Let $X_1, \ldots, X_n$ *iid* Bernoulli with parameter p.
- $E(X_i) = p$.
- $X = \sum_{i=1}^{n} X_i \sim Bin(n, p)$.
- $E(X) = np$.

Example (Expectation of a negative binomial random variable)

- Let $X_1, \ldots, X_n$ *iid* geometric with parameter p.
- $E(X_i) = \frac{1}{p}$.
- $X = \sum_{i=1}^{r} X_i \sim NB(r, p)$.
- $E(X) = \frac{r}{p}$.

Expectation of Sums of Random Variables

Example (Expectation of a hypergeometric random variable)

If n balls are randomly selected from an urn containing N balls of which m are white, find the expected number of white balls selected.

- Let
$$X_i = \begin{cases} 1 & \text{if the } i^{th} \text{ white ball is selected} \\ 0 & \text{otherwise} \end{cases},$$
- $E(X_i) = P(X_i = 1)$,
$$\begin{aligned} E(X_i) &= P(i^{th} \text{ white ball is selected}) \\ &= \frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}} \\ &= \frac{n}{N} \end{aligned}$$
- $X = \sum_{i=1}^{m} X_i$ denote the number of white balls selected.
- $E(X) = \frac{mn}{N}$.

Expectation of Sums of Random Variables

- When one is dealing with an infinite collection of random variables X_i, $i \geq 1$, each having a finite expectation ($E(X_i) < \infty$), it is not necessarily true that

$$E\left(\sum_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} E(X_i).$$

- To determine when it is valid, we note that $\sum_{i=1}^{\infty} X_i = \lim_{n\to\infty} \sum_{i=1}^{n} X_i$. Thus,

$$E\left(\sum_{i=1}^{\infty} X_i\right) = E\left(\lim_{n\to\infty} \sum_{i=1}^{n} X_i\right) \stackrel{?}{=} \lim_{n\to\infty} E\left(\sum_{i=1}^{n} X_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} E\left(X_i\right) = \sum_{i=1}^{\infty} E(X_i)$$

- It is valid whenever we are justified in **interchanging the expectation and limit operations**.
- Although, in general, this interchange is not justified, it can be shown to be valid in two important special cases:
  1. The X_i are all nonnegative random variables. (That is, $P(X_i \geq 0) = 1$ for all i).
  2. $\sum_{i=1}^{\infty} E(|X_i|) < \infty$.

Moments of the Number of Events that Occur

- For given events $A_1, A_2, \ldots, A_n$, let X be the number of these events that occur.
- Define

$$I_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases},$$

- Because $X = \sum_{i=1}^{n} I_i$, we obtain $E(X) = \sum_{i=1}^{n} P(A_i)$.
- Now suppose we are interested in the number of pairs of events that occur.
- Then,

$$I_i I_j = \begin{cases} 1 & \text{if both } A_i \& A_j \text{ occur} \\ 0 & \text{otherwise} \end{cases},$$

- It follows that the number of pairs is equal to $\sum_{i<j} I_i I_j$.
- But because X is the number of events that occur, it also follows that the number of pairs of events that occur is $\binom{X}{2}$.

Moments of the Number of Events that Occur

- $\binom{X}{2} = \sum_{i<j} I_i I_j$, where there are $\binom{n}{2}$ terms in the summation.
- Taking expectations yields

$$E\left(\binom{X}{2}\right) = \sum_{i<j} E\left(I_i I_j\right) = \sum_{i<j} P\left(A_i \cap A_j\right)$$

- or

$$E\left(\frac{X(X-1)}{2}\right) = \sum_{i<j} P\left(A_i \cap A_j\right)$$

- giving that $E(X^2) - E(X) = 2\sum_{i<j} P\left(A_i \cap A_j\right)$.
- which yields $E(X^2)$, and thus, $Var(X) = E(X^2) - (E(X))^2$.
- In general,

$$E\left(\binom{X}{k}\right) = \sum_{i_1<i_2<\cdots<i_k} E\left(I_{i_1} I_{i_2} \cdots I_{i_k}\right) = \sum_{i_1<i_2<\cdots<i_k} P\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right)$$

Moments of the Number of Events that Occur

Example (Moments of binomial random variables)

- Consider n independent trials, with each trial being a success with probability p.
- Let A_i be the event that trial i is a success.
- When $i \neq j$, $P(A_i \cap A_j) = p^2$.
- Consequently,

$$E\left(\binom{X}{2}\right) = \sum_{i<j} p^2 = \binom{n}{2} p^2.$$

- or

$$E(X^2) - E(X) = n(n-1)p^2.$$

- But,

$$E(X) = \sum_{i=1}^{n} P(A_i) = np.$$

- From the preceding slide,

$$Var(X) = n(n-1)p^2 + np - (np)^2 = np(1-p).$$

Theorem

If X and Y are independent, then, for any functions h and g,

$$E\left(g(X)h(Y)\right) = E\left(g(X)\right)E\left(h(Y)\right).$$

Proof:

Definition (Covariance)

The covariance between X and Y, denoted by $Cov(X, Y)$, is defined by

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

- $Cov(X, Y)$ gives us information about the relationship between the random variables X and Y.

- It can be shown that $Cov(X, Y) = E(XY) - E(X)E(Y)$.
  Proof:

Covariance, Variance of Sums, and Correlations

- If X and Y are independent, then $Cov(X, Y) = 0$.
- However, the converse is not true.

  Counter example:
  Let X be a random variable such that

  $$P(X = 0) = P(X = 1) = P(X = -1) = \frac{1}{3},$$

  and define Y:

  $$Y = \begin{cases} 1 & \text{if } X = 0 \\ 0 & \text{if } X \neq 0 \end{cases},$$

  - $XY = 0 \implies E(XY) = 0$.
  - Also, $E(X) = 0$.
  - Thus, $Cov(X, Y) = 0$.
  - However, it is clear that X and Y are dependent.

Covariance, Variance of Sums, and Correlations

Theorem

1. $Cov(X, Y) = Cov(Y, X)$.

2. $Cov(X, X) = Var(X)$.

3. $Cov(aX, Y) = aCov(X, Y)$.

4. $Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$.

Covariance, Variance of Sums, and Correlations

Theorem

$$\begin{aligned}
Var\left(\sum_{i=1}^{n} X_i\right) &= Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} Var(X_i) + \sum_{i\neq j} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)
\end{aligned}$$

Theorem

If $X_1, \ldots, X_n$ are pairwise independent, in that X_i and X_j are independent for $i \neq j$, then

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i).$$

Covariance, Variance of Sums, and Correlations

- $Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i<j} a_i a_j Cov(X_i, X_j).$

- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y).$

- $Cov(aX + b, cY + d) = acCov(X, Y).$

Covariance, Variance of Sums, and Correlations

Example

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having expected value μ and variance σ^2.

- Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean.
- The quantities $X_i - \bar{X}$, $i = 1, \ldots, n$, are called **deviations**, as they equal the differences between the individual data and the sample mean.
- The random variable $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is called the sample variance.

(1) Find $Var(\bar{X})$.

(2) Show that $E(S^2) = \sigma^2$.

Covariance, Variance of Sums, and Correlations

Definition (Correlation)

The correlation of two random variables X and Y, denoted by $\rho(X, Y)$, is defined, as long as $Var(X)Var(Y)$ is positive, by

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

Theorem

$$-1 \leq \rho(X, Y) \leq 1.$$

Proof:

- *Hint: Use Cauchy-Schwarz inequality:*

$$\left(\sum_{i=1}^{n} X_i^{'} Y_i^{'}\right)^2 \leq \left(\sum_{i=1}^{n} X_i^{'}\right)\left(\sum_{i=1}^{n} Y_i^{'}\right).$$

- *Set* $X_i^{'} = (X_i - \bar{X})$ *and* $Y_i^{'} = (Y_i - \bar{Y})$.

Covariance, Variance of Sums, and Correlations

Properties of ρ:

1. $-1 \leq \rho(X, Y) \leq 1$. (Previous Slide)
2. $\rho(X, Y) = 1$ if and only if $Y = aX + b$ for some $a > 0$.
3. $\rho(X, Y) = -1$ if and only if $Y = aX + b$ for some $a < 0$.
4. $\rho(X, Y)$ measures the degree of linearity between X and Y. A value near $+1$ or -1 indicates a high degree of linearity, whereas a value near 0 indicates that such linearity is absent.
5. A positive value indicates that Y tends to increase when X does, whereas a negative value indicates that Y tends to decrease when X increases.
6. $\rho(X, Y) = 0$, then X and Y are said to be uncorrelated.

Example

Suppose X and Y have a joint *pdf*:

$$f(x,y) = \begin{cases} 2e^{-x}e^{-y} & \text{if } 0 \le y \le x < \infty \\ 0 & \text{otherwise} \end{cases},$$

(1) The marginal *pdf* of X is given by

$$f_X(x) = \int_0^x 2e^{-x}e^{-y}dy = 2e^{-x}\left(1 - e^{-x}\right), \quad x \ge 0.$$

(2) The marginal *pdf* of Y is given by

$$f_Y(y) = \int_y^\infty 2e^{-x}e^{-y}dx = 2e^{-2y}, \quad y \ge 0.$$

Covariance, Variance of Sums, and Correlations

(3) $E(X) = \frac{3}{2}$, and $Var(X) = \frac{5}{4}$.

(4) $E(Y) = \frac{1}{2}$, and $Var(Y) = \frac{1}{4}$.

(5) $E(XY) = 1$.

(6) $Cov(X, Y) = 1 - \frac{3}{2} \times \frac{1}{2} = \frac{1}{4}$.

(7) $Cor(X, Y) = \rho(X, Y) = \frac{1}{\sqrt{5}}$.

Covariance, Variance of Sums, and Correlations

Example

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having variance σ^2. Show that

$$Cov\left(X_i - \bar{X}, \bar{X}\right) = 0.$$

- We have

$$\begin{aligned}
Cov\left(X_i - \bar{X}, \bar{X}\right) &= Cov\left(X_i, \bar{X}\right) - Cov\left(\bar{X}, \bar{X}\right) \\
&= Cov\left(X_i, \frac{1}{n}\sum_{j=1}^{n} X_j\right) - Var(\bar{X}) \\
&= \frac{1}{n}\sum_{j=1}^{n} Cov(X_i, X_j) - \frac{\sigma^2}{n} \\
&= 0.
\end{aligned}$$

where

$$Cov(X_i, X_j) = \begin{cases} 0 & \text{if } i \neq j \text{ by independence} \\ \sigma^2 & \text{if } i = j \text{ since } Var(X_i) = \sigma^2 \end{cases},$$

Covariance, Variance of Sums, and Correlations

- Although $\bar{X}$ and the deviation $X_i - \bar{X}$ are uncorrelated, they are not, in general, independent.

- However, in the special case where the X_i are normal random variables, it turns out that not only is $\bar{X}$ independent of a single deviation, but it is independent of the entire sequence of deviations $X_j - \bar{X}$, $j = 1, \ldots, n$.

- The sample mean $\bar{X}$ and the sample variance S^2 are independent.

- $\frac{(n-1)S^2}{\sigma^2}$ have a chi-squared distribution with $n - 1$ degrees of freedom (More details later).

Conditional Expectation

Definition (Conditional Expectation – Discrete Case)

- Recall that if X and Y are jointly discrete random variables, then the conditional probability mass function of X, given that $Y = y$, is defined for all y such that $P(Y = y) > 0$, by

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)}.$$

- It is therefore natural to define, in this case, the conditional expectation of X given that $Y = y$, for all values of y such that $p_Y(y) > 0$, by

$$\begin{aligned} E(X|Y = y) &= \sum_x xP(X = x|Y = y) \\ &= \sum_x xp_{X|Y}(x|y). \end{aligned}$$

- Therefore, if X and Y are independent, then $E(X|Y = y) = E(X)$.

Conditional Expectation

Example

If X and Y are independent binomial random variables with identical parameters n and p, calculate the conditional expected value of X given that $X+Y=m$.

- Let us first calculate the conditional *pmf* of X given that $X+Y=m$.

$$\begin{aligned} P(X=k|X+Y=m) &= \frac{P(X=k, X+Y=m)}{P(X+Y=m)} \\ &= \frac{P(X=k, Y=m-k)}{P(X+Y=m)} \\ &= \frac{P(X=k)\,P(Y=m-k)}{P(X+Y=m)} \\ &= \frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}}, \quad k \le \min(m,n). \end{aligned}$$

- Hence, the conditional distribution of X, given that $X+Y=m$, is the hypergeometric distribution.
- Therefore,

$$E(X|X+Y=m) = \frac{mn}{2n} = \frac{m}{2}.$$

Conditional Expectation

Definition (Conditional Expectation – Continuous Case)

- If X and Y are jointly continuous random variables with a joint probability density function $f(x, y)$, then the conditional probability density of X, given that $Y = y$, is defined for all y such that $f_y(y) > 0$, by

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

- It is therefore natural to define, in this case, the conditional expectation of X given that $Y = y$, for all values of y such that $f_Y(y) > 0$, by

$$E\left(X|Y = y\right) = \int_{-\infty}^{\infty} x f_{X|Y}\left(x|y\right) dx.$$

Conditional Expectation

Example

Suppose that the joint density of X and Y is given by

$$f(x,y) = \frac{e^{-x/y}e^{-y}}{y}, \quad 0 < x < \infty,\ 0 < y < \infty.$$

Compute $E\left(X|Y = y\right)$.

Ans. $X|Y \sim Exp(1/y)$.

Conditional Expectation

- Just as conditional probabilities satisfy all of the properties of ordinary probabilities, so do conditional expectations satisfy the properties of ordinary expectations.

- For instance, such formulas as

$$E\left(g(X)|Y=y\right)=\begin{cases}\sum_x g(x)p_{X|Y}(x|y) & \text{in the discrete case}\\ \int\limits_{-\infty}^{\infty} g(x)f_{X|Y}(x|y)dx & \text{in the continuous case}\end{cases},$$

- and

$$E\left(\sum_{i=1}^{n} X_i|Y=y\right)=\sum_{i=1}^{n} E\left(X_i|Y=y\right).$$

Conditional Expectation

Theorem

Let us denote by $E(X|Y)$ that function of the random variable Y whose value at $Y = y$ is $E(X|Y = y)$.

$$E(X) = E\left[E\left(X|Y\right)\right] = \begin{cases} \sum_y E\left(X|Y = y\right) P(Y = y) & \textit{If } Y \textit{ is discrete} \\ \int\limits_{-\infty}^{\infty} E\left(X|Y = y\right) f_Y(y) dy & \textit{If } Y \textit{ is continuous} \end{cases},$$

Proof:

- See Example 5f Page 319.

Conditional Expectation

Definition (Conditional Variance)

The conditional variance of X given that $Y = y$ is defined by

$$Var(X|Y) = E\left[(X - E(X|Y))^2 \,|Y\right]$$

- It can be written as

$$Var(X|Y) = E(X^2|Y) - (E(X|Y))^2 \,.$$

- $Var(X|Y)$ is exactly analogous to the usual definition of variance, but now all expectations are conditional on the fact that Y is known.

Conditional Expectation

Theorem

$$Var(X) = E\left[Var(X|Y)\right] + Var\left[E(X|Y)\right].$$

Proof:

- $Var(X|Y) = E\left(X^2|Y\right) - (E(X|Y))^2.$
- $E\left[Var(X|Y)\right] = E(X^2) - E\left[(E(X|Y))^2\right].$
- $Var\left[E\left(X|Y\right)\right] = E\left[(E(X|Y))^2\right] - (E(X))^2.$

Moment Generating Functions

Definition (Moment Generating Function)

The moment generating function $M(t)$ of the random variable X is defined for all real values of t by

$$M(t) = E\left(e^{tX}\right) = \begin{cases} \sum_x e^{tx} p(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{if } X \text{ is continuous} \end{cases},$$

- We call $M(t)$ the moment generating function because all the moments of X can be obtained by successively differentiating $M(t)$ and then evaluating the result at $t = 0$.
- For example,

$$M^{'}(t) = \frac{d}{dt} E\left(e^{tX}\right) = E\left(\frac{d}{dt}\left(e^{tX}\right)\right) = E\left(Xe^{tX}\right)$$

  where we have assumed that the interchange of the differentiation and expectation operators is legitimate.

Moment Generating Functions

- Hence,

$$M^{'}(0) = E(X).$$

- Similarly,

$$M^{''}(t) = E\left(X^2 e^{tX}\right).$$

- Thus,

$$M^{''}(0) = E(X^2).$$

- In general, the n^{th} derivative of $M(t)$ is given by

$$M^{(n)}(t) = E\left(X^n e^{tX}\right), \quad n \geq 1.$$

- Implying that,

$$M^{(n)}(0) = E\left(X^n\right), \quad n \geq 1.$$

Moment Generating Functions

Example

If X is a binomial random variable with parameters n and p, then

$$M(t) = E\left(e^{tX}\right) = \left(pe^{t} + (1-p)\right)^{n}$$

Moment Generating Functions

Example

If X is a Poisson random variable with parameter λ, then

$$M(t) = E\left(e^{tX}\right) = e^{\lambda\left(e^{t}-1\right)}$$

Moment Generating Functions

Example

If X is an exponential random variable with parameter λ, then

$$M(t) = E\left(e^{tX}\right) = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

We note from this derivation that, for the exponential distribution, $M(t)$ defined only for values of $t < \lambda$.

Moment Generating Functions

Example

Let X be a normal random variable with parameters μ and σ^2.

- We first compute the moment generating function of a standard normal random variable with parameters 0 and 1.
- Letting Z be such a random variable, we have

$$\begin{aligned}
M(t) &= E\left(e^{tX}\right) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{2}x^2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2-2tx)} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-t)^2+\frac{1}{2}t^2} dx \\
&= e^{\frac{1}{2}t^2}.
\end{aligned}$$

Moment Generating Functions

Example

- Hence, the moment generating function of the standard normal random variable Z is $M_Z(t) = e^{\frac{1}{2}t^2}$.
- We now compute the moment generating function of a normal random variable $X = \mu + \sigma Z$ with parameters μ and σ^2.
- We have
$$M(t) = E\left(e^{tX}\right) = E\left(e^{t(\mu+\sigma Z)}\right) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}.$$

Moment Generating Functions

Discrete Probability Distribution.				
	Probability mass function, $p(x)$	Moment generating function, $M(t)$	Mean	Variance
Binomial with parameters n, p; $0 \leq p \leq 1$	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \ldots, n$	$(pe^t + 1 - p)^n$	np	$np(1-p)$
Poisson with parameter $\lambda > 0$	$e^{-\lambda} \frac{\lambda^x}{x!}$ $x = 0, 1, 2, \ldots$	$\exp\{\lambda(e^t - 1)\}$	λ	λ
Geometric with parameter $0 \leq p \leq 1$	$p(1-p)^{x-1}$ $x = 1, 2, \ldots$	$\frac{pe^t}{1-(1-p)e^t}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negative binomial with parameters r, p; $0 \leq p \leq 1$	$\binom{n-1}{r-1} p^r (1-p)^{n-r}$ $n = r, r+1, \ldots$	$\left[\frac{pe^t}{1-(1-p)e^t}\right]^r$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$

Moment Generating Functions

Continuous Probability Distribution.				
	Probability density function, $f(x)$	Moment generating function, $M(t)$	Mean	Variance
Uniform over (a, b)	$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential with parameter $\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\frac{\lambda}{\lambda - t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma with parameters $(s, \lambda), \lambda > 0$	$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{s-1}}{\Gamma(s)} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\left(\frac{\lambda}{\lambda - t}\right)^s$	$\frac{s}{\lambda}$	$\frac{s}{\lambda^2}$
Normal with parameters (μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$	$\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$	μ	σ^2

Moment Generating Functions

Theorem

The moment generating function uniquely determines the distribution*. That is, if $M_X(t)$ exists and is finite in some region about $t = 0$, then the distribution of X is uniquely determined.*

Theorem

The moment generating function of the sum of independent random variables equals the product of the individual moment generating functions.

Proof:

Moment Generating Functions

Example

Let X and Y be independent binomial random variables with parameters (n, p) and (m, p), respectively. What is the distribution of $X + Y$?

Moment Generating Functions

Example

Let X and Y be independent Poisson random variables with parameters λ_1 and λ_2, respectively. What is the distribution of $X + Y$?

Moment Generating Functions

Example

Let X and Y be independent normal random variables with respective parameters (μ_1, σ_1^2) and (μ_2, σ_2^2). What is the distribution of $X + Y$?

Ans. $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Moment Generating Functions

Definition (Joint Moment Generating Functions)

For any n random variables $X_1, \dots, X_n$, the **joint moment generating function**, $M(t_1, \dots, t_n)$, is defined, for all real values of $t_1, \dots, t_n$, by

$$M(t_1, \dots, t_n) = E\left(e^{t_1 X_1 + \cdots + t_n X_n}\right)$$

- The individual moment generating functions can be obtained from $M(t_1, \dots, t_n)$ by letting all but one of the t_j's be 0.

$$M_{X_i}(t) = E\left(e^{tX_i}\right) = M(0, \dots, 0, i, 0, \dots, 0)$$

  where the t is in the i^{th} place.

- The joint moment generating function $M(t_1, \dots, t_n)$ uniquely determines the joint distribution of $X_1, \dots, X_n$.

Moment Generating Functions

Theorem

$X_1, \ldots, X_n$ are independent random variables if and only if

$$M(t_1, \ldots, t_n) = M_{X_1}(t_1) \times \cdots \times M_{X_n}(t_n).$$

Proof:

Moment Generating Functions

Example

Let X and Y be independent normal random variables, each with mean μ and variance σ^2. Show that $X + Y$ and $X - Y$ are independent.

Problems and Exercises

PROBLEMS

PAGE 352:

4, 6, 7, 30, 31, 33, 37, 38, 40, 45, 48, 50, 51

THEORETICAL EXERCISES

PAGE 359:

1, 2, 19, 20, 22, 25, 26, 28