Math 461 Spring 2024

Renming Song

University of Illinois Urbana-Champaign

April 05, 2024

Outline

Outline

HW9 is due today, before the end of class.

Test 2 is next Friday. Materials covered on Test 2: Section 4.9, Section 5.1, Section 5.2, Section 5.3, Section 5.4, Section 5.5, Section 5.6, Section 5.7, Section 6.1, Section 6.2, Section 6.3, Section 6.4, Section 6.5, Section 6.6, Section 7.2, Section 7.4

HW9 is due today, before the end of class.

Test 2 is next Friday. Materials covered on Test 2: Section 4.9, Section 5.1, Section 5.2, Section 5.3, Section 5.4, Section 5.5, Section 5.6, Section 5.7, Section 6.1, Section 6.2, Section 6.3, Section 6.4, Section 6.5, Section 6.6, Section 7.2, Section 7.4

Outline

Example 7

Independent trials, each resulting in a success with probability p, are performed until a success occurs. Let N be the number of trials needed. N is a geometric random variable with parameter p. Find $E[N]$ and $\mathrm{Var}(N)$.

Let $Y = 1$ if the first trial results in a success and $Y = 0$ otherwise. Then

$$\begin{aligned} E[N] &= E[E[N|Y]] = E[N|Y=1]P(Y=1) + E[N|Y=0]P(Y=0) \\ &= 1 \cdot p + (1 + E[N])(1-p) \end{aligned}$$

which implies $pE[N] = 1$, that is, $E[N] = \frac{1}{p}$.

Example 7

Independent trials, each resulting in a success with probability p, are performed until a success occurs. Let N be the number of trials needed. N is a geometric random variable with parameter p. Find $E[N]$ and $\mathrm{Var}(N)$.

Let $Y = 1$ if the first trial results in a success and $Y = 0$ otherwise. Then

$$\begin{aligned} E[N] = E[E[N|Y]] &= E[N|Y = 1]P(Y = 1) + E[N|Y = 0]P(Y = 0) \\ &= 1 \cdot p + (1 + E[N])(1 - p) \end{aligned}$$

which implies $pE[N] = 1$, that is, $E[N] = \frac{1}{p}$.

$$E[N^2] = E[E[N^2|Y]] = E[N^2|Y=1]P(Y=1) + E[N^2|Y=0]P(Y=0)$$
$$= 1 \cdot p + E[(1+N)^2](1-p)$$
$$= 1 + (1-p)E[2N + N^2] = 1 + \frac{2(1-p)}{p} + (1-p)E[N^2].$$

Solving for $E[N^2]$, we get

$$E[N^2] = \frac{2-p}{p^2}.$$

Thus

$$\mathrm{Var}(N) = E[N^2] - (E[N])^2 = \frac{1-p}{p^2}.$$

$$
\begin{aligned}
E[N^2] &= E[E[N^2|Y]] = E[N^2|Y=1]P(Y=1) + E[N^2|Y=0]P(Y=0) \\
&= 1 \cdot p + E[(1+N)^2](1-p) \\
&= 1 + (1-p)E[2N + N^2] = 1 + \frac{2(1-p)}{p} + (1-p)E[N^2].
\end{aligned}
$$

Solving for $E[N^2]$, we get

$$E[N^2] = \frac{2-p}{p^2}.$$

Thus

$$\mathrm{Var}(N) = E[N^2] - (E[N])^2 = \frac{1-p}{p^2}.$$

Example 8

A coin, having probability $p \in (0, 1)$ of landing heads, is continually flipped until at least one head and one tail have been flipped.

(a) Find the expected number of flips needed, and its variance.

(b) Find the expected number of flips that land on heads.

(a) Let X be the number of flips needed. Let $Y = 1$ if the first flip is H, and $Y = 0$ the first flip is T. Then

$$E[X] = E\left[E\left[X|Y\right]\right].$$

Given the first flip is H, the number of additional flips needed $(X - 1)$ is a geometric random variable with parameter $1 - p$. Thus

$$E\left[X|Y = 1\right] = 1 + \frac{1}{1-p}.$$

Example 8

A coin, having probability $p \in (0, 1)$ of landing heads, is continually flipped until at least one head and one tail have been flipped.

(a) Find the expected number of flips needed, and its variance.

(b) Find the expected number of flips that land on heads.

(a) Let X be the number of flips needed. Let $Y = 1$ if the first flip is H, and $Y = 0$ the first flip is T. Then

$$E[X] = E\left[E\left[X|Y\right]\right].$$

Given the first flip is H, the number of additional flips needed $(X - 1)$ is a geometric random variable with parameter $1 - p$. Thus

$$E\left[X|Y = 1\right] = 1 + \frac{1}{1-p}.$$

Similarly, given the first flip is T, the number of additional flips needed $(X-1)$ is a geometric random variable with parameter p, and

$$E\left[X|Y=0\right] = 1+\frac{1}{p}.$$

Hence

$$\begin{aligned}E[X] &= E\left[E\left[X|Y\right]\right] = p\cdot E\left[X|Y=1\right]+(1-p)\cdot E\left[X|Y=0\right]\\ &= p\cdot\left(1+\frac{1}{1-p}\right)+(1-p)\cdot\left(1+\frac{1}{p}\right) = 1+\frac{p}{1-p}+\frac{1-p}{p}.\end{aligned}$$

Given $Y=1$, $X-1$ is a geometric random variable with parameter $1-p$, thus

$$E\left[X^2|Y=1\right] = 1+\frac{2}{1-p}+\left(\frac{p}{(1-p)^2}+\frac{1}{(1-p)^2}\right) = 1+\frac{3-p}{(1-p)^2}.$$

Given $Y = 0$, $X - 1$ is a geometric random variable with parameter p, thus

$$E\left[X^2 | Y = 0\right] = 1 + \frac{2}{p} + \left(\frac{1-p}{p^2} + \frac{1}{p^2}\right) = 1 + \frac{2+p}{p^2}.$$

Hence

$$\begin{aligned} E[X^2] &= E\left[E\left[X^2|Y\right]\right] = p \cdot E\left[X^2 | Y = 1\right] + (1-p) \cdot E\left[X^2 | Y = 0\right] \\ &= p \cdot \left(1 + \frac{3-p}{(1-p)^2}\right) + (1-p) \cdot \left(1 + \frac{2+p}{p^2}\right) \\ &= 1 + \frac{p(3-p)}{(1-p)^2} + \frac{(1-p)(2+p)}{p^2}. \end{aligned}$$

Therefore

$$\begin{aligned} \operatorname{Var}[X] &= E\left[X^2\right] - (E[X])^2 \\ &= 1 + \frac{p(3-p)}{(1-p)^2} + \frac{(1-p)(2+p)}{p^2} - \left(1 + \frac{p}{1-p} + \frac{1-p}{p}\right)^2. \end{aligned}$$

Given $Y = 0$, $X - 1$ is a geometric random variable with parameter p, thus

$$E\left[X^2|Y=0\right] = 1 + \frac{2}{p} + \left(\frac{1-p}{p^2} + \frac{1}{p^2}\right) = 1 + \frac{2+p}{p^2}.$$

Hence

$$\begin{aligned} E[X^2] &= E\left[E\left[X^2|Y\right]\right] = p \cdot E\left[X^2|Y=1\right] + (1-p) \cdot E\left[X^2|Y=0\right] \\ &= p \cdot \left(1 + \frac{3-p}{(1-p)^2}\right) + (1-p) \cdot \left(1 + \frac{2+p}{p^2}\right) \\ &= 1 + \frac{p(3-p)}{(1-p)^2} + \frac{(1-p)(2+p)}{p^2}. \end{aligned}$$

Therefore

$$\begin{aligned} \mathrm{Var}[X] &= E\left[X^2\right] - (E[X])^2 \\ &= 1 + \frac{p(3-p)}{(1-p)^2} + \frac{(1-p)(2+p)}{p^2} - \left(1 + \frac{p}{1-p} + \frac{1-p}{p}\right)^2. \end{aligned}$$

(b) Let Z be the number of flips that lands on heads. Then, given $Y = 1$, the number of flips that land on heads is a geometric random variable with parameter $1 - p$, thus

$$E[Z|Y = 1] = \frac{1}{1-p}.$$

Given $Y = 0$, the number of flips that land on heads is equal to 1, thus

$$E[Z|Y = 0] = 1.$$

Consequently

$$E[Z] = p \cdot \frac{1}{1-p} + (1-p) \cdot 1 = \frac{p}{1-p} + 1 - p.$$

Outline

The moment generating function of a random variable X is defined to be the function

$$M_X(t) = E[e^{tX}], \quad t \in \mathbb{R}.$$

$M_X(t)$ may not be defined for all $t \in \mathbb{R}$, but it is always defined for $t = 0$. In fact, $M_X(0) = 1$. We will concentrate on random variables X for which $M_X(t)$ is defined at least in an interval around the origin. All the important random variables we learned in the course satisfy this property.

Why the name "moment generating function"? For a random variable X satisfying the property above, one can justify that for t in that interval,

The moment generating function of a random variable X is defined to be the function

$$M_X(t) = E[e^{tX}], \quad t \in \mathbb{R}.$$

$M_X(t)$ may not be defined for all $t \in \mathbb{R}$, but it is always defined for $t = 0$. In fact, $M_X(0) = 1$. We will concentrate on random variables X for which $M_X(t)$ is defined at least in an interval around the origin. All the important random variables we learned in the course satisfy this property.

Why the name "moment generating function"? For a random variable X satisfying the property above, one can justify that for t in that interval,

The moment generating function of a random variable X is defined to be the function

$$M_X(t) = E[e^{tX}], \quad t \in \mathbb{R}.$$

$M_X(t)$ may not be defined for all $t \in \mathbb{R}$, but it is always defined for $t = 0$. In fact, $M_X(0) = 1$. We will concentrate on random variables X for which $M_X(t)$ is defined at least in an interval around the origin. All the important random variables we learned in the course satisfy this property.

Why the name "moment generating function"? For a random variable X satisfying the property above, one can justify that for t in that interval,

$$M_X'(t) = E[Xe^{tX}], M_X''(t) = E[X^2e^{tX}], \ldots, M_X^{(n)}(t) = E[X^ne^{tX}].$$

Thus

$$M_X'(0) = E[X], M_X''(0) = E[X^2], \ldots M_X^{(n)}(0) = E[X^n].$$

and

$$\mathrm{Var}(X) = M_X''(0) - (M_X'(0))^2.$$

Once we know the moment generating function $M_X(t)$ of X, we can easily find all the moments of X. This is why we call $M_X(t)$ he moment generating function of X

$$M_X'(t) = E[Xe^{tX}], M_X''(t) = E[X^2e^{tX}], \ldots, M_X^{(n)}(t) = E[X^ne^{tX}].$$

Thus

$$M_X'(0) = E[X], M_X''(0) = E[X^2], \ldots M_X^{(n)}(0) = E[X^n].$$

and

$$\mathrm{Var}(X) = M_X''(0) - (M_X'(0))^2.$$

Once we know the moment generating function $M_X(t)$ of X, we can easily find all the moments of X. This is why we call $M_X(t)$ he moment generating function of X

Proposition

(i) If X is a binomial random variable with parameters (n, p), then

$$M_X(t) = (pe^t + 1 - p)^n, \quad \text{for all } t \in \mathbb{R}.$$

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

$$M_X(t) = e^{\lambda(e^t - 1)}, \quad \text{for all } t \in \mathbb{R}.$$

(iii) If X is a geometric random variable with parameter p, then

$$M_X(t) = \frac{pe^t}{1 - (1-p)e^t}, \quad \text{for all } t < -\ln(1-p).$$

(iv) If X is a negative binomial random variable with parameters (r, p), then

$$M_X(t) = \left(\frac{pe^t}{1 - (1-p)e^t} \right)^r, \quad \text{for all } t < -\ln(1-p).$$

Proposition (cont)

(v) If X is uniformly distributed in the interval (a, b), then

$$M_X(t) = \frac{e^{tb} - e^{ta}}{b - a}, \quad \text{for all } t \in \mathbb{R}.$$

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

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad \text{for all } t < \lambda.$$

(vii) If X is a Gamma random variable with parameters (α, λ), then

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}, \quad \text{for all } t < \lambda.$$

(viii) If X is a normal random variable with parameters (μ, σ^2), then

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right), \quad \text{for all } t \in \mathbb{R}.$$

Let's derive two of the 8 items above. Suppose X is a Poisson random variable with parameter λ, then

$$M_X(t) = E[e^{tX}] = \sum_{i=0}^{\infty} e^{ti} e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda e^t)^i}{i!}$$
$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}.$$

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

$$M_X(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx$$
$$= \frac{\lambda}{\lambda - t}, \quad \text{for all } t < \lambda.$$

Let's derive two of the 8 items above. Suppose X is a Poisson random variable with parameter λ, then

$$M_X(t) = E[e^{tX}] = \sum_{i=0}^{\infty} e^{ti} e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda e^t)^i}{i!}$$
$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}.$$

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

$$M_X(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx$$
$$= \frac{\lambda}{\lambda - t}, \quad \text{for all } t < \lambda.$$

Theorem

If X and Y are two random variables with $M_X(t) = M_Y(t)$ for all t, then X and Y have the same distribution.

This theorem says that the moment generating function $M_X(t)$ of X also contains all the statistical information about X.

Theorem

If X and Y are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t), \qquad \text{for all } t.$$

Theorem

If X and Y are two random variables with $M_X(t) = M_Y(t)$ for all t, then X and Y have the same distribution.

This theorem says that the moment generating function $M_X(t)$ of X also contains all the statistical information about X.

Theorem

If X and Y are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t), \qquad \text{for all } t.$$

Theorem

If X and Y are two random variables with $M_X(t) = M_Y(t)$ for all t, then X and Y have the same distribution.

This theorem says that the moment generating function $M_X(t)$ of X also contains all the statistical information about X.

Theorem

If X and Y are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t), \quad \text{for all } t.$$

Since X and Y are independent, e^{tX} and e^{tY} are also independent. Thus

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}] = M_X(t)M_Y(t).$$

Next time, I will use this theorem to prove that sums of independent binomial random variables with a common second parameter p is again a binomial random variable, and other similar results.

Since X and Y are independent, e^{tX} and e^{tY} are also independent. Thus

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}] = M_X(t)M_Y(t).$$

Next time, I will use this theorem to prove that sums of independent binomial random variables with a common second parameter p is again a binomial random variable, and other similar results.