Lecture 6: Expected Value and Moments

Sta 111

Colin Rundel

May 21, 2014

Expected Value

The expected value of a random variable is defined as follows

Discrete Random Variable:

$$E[X] = \sum_{\text{all } x} xP(X = x)$$

Continous Random Variable:

$$E[X] = \int_{\text{all } x} xP(X = x)dx$$

Expected Value of a function

The expected value of a function of a random variable is defined as follows

Discrete Random Variable:

$$E[f(X)] = \sum_{\text{all } x} f(x)P(X = x)$$

Continous Random Variable:

$$E[f(X)] = \int_{\text{all } x} f(x)P(X = x)dx$$

Properties of Expected Value

- **Constants** - $E(c) = c$ if c is constant
- **Indicators** - $E(I_A) = P(A)$ where I_A is an indicator function
- **Constant Factors** - $E(cX) = cE(x)$
- **Addition** - $E(X + Y) = E(X) + E(Y)$
- **Multiplication** - $E(XY) = E(X)E(Y)$ if X and Y are independent.

Variance

Another common property of random variables we are interested in is the Variance which measures the squared deviation from the mean.

$$Var(X) = E\left[(X - E(X))^2\right] = E(X - \mu)^2$$

One common simplification:

$$\begin{aligned} Var(X) &= E(X - \mu)^2 \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

Standard Deviation:

$$SD(X) = \sqrt{Var(X)}$$

Properties of Variance

What is $Var(aX + b)$ when a and b are constants?

Which gives us:

$$\begin{aligned} Var(aX) &= a^2 Var(X) \\ Var(X + c) &= Var(X) \\ Var(c) &= 0 \end{aligned}$$

Properties of Variance, cont.

What about $Var(X + Y)$?

Covariance

What about when X and Y are not independent?

$$E(XY) \neq E(X)E(Y) \quad \Rightarrow \quad E(XY) - \mu_x\mu_y \neq 0$$

This quantity is known as Covariance, and is roughly speaking a generalization of variance to two variables

$$\begin{aligned} Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[(X - \mu_x)(Y - \mu_y)] \\ &= E[XY + \mu_x\mu_y - X\mu_y - Y\mu_x]) \\ &= E(XY) - \mu_x\mu_y \end{aligned}$$

Properties of Covariance

- $Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - \mu_x\mu_y$
- $Cov(X, Y) = 0$ if X and Y are independent
- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$
- $Cov(aX, bY) = ab\ Cov(X, Y)$
- $Cov(X + a, Y + b) = Cov(X, Y)$

Properties of Variance, cont.

A general formula for the variance of the linear combination of two random variables:

From which we can see that

$$Var(X + Y) = Var(X) + Var(Y) + Cov(X, Y)$$
$$Var(X - Y) = Var(X) + Var(Y) - Cov(X, Y)$$

Properties of Variance, cont.

For a completely general formula for the variances of a linear combination of n random variables:

$$Var\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(c_i X_i, c_j X_j)$$
$$= \sum_{i=1}^{n} c_i^2\ Var(X_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} c_i c_j\ Cov(X_i, X_j)$$

Bernoulli Random Variable

Let $X \sim \text{Bern}(p)$, what is $E(X)$ and $Var(X)$?

Binomial Random Variable

Let $X \sim \text{Binom}(n, p)$, what is $E(X)$ and $Var(X)$?

We can redefine $X = \sum_{i=1}^{n} Y_i$ where $Y_1, \cdots, Y_n \sim \text{Bern}(p)$, and since we are *sampling with replacement* all Y_i and Y_j are independent.

Hypergeometric Random Variable - $E(X)$

Lets consider a simple case where we have an urn with m black marbles and $N - m$ white marbles. Let B_i be an indicator variable for the ith marble being black.

$$B_i = \begin{cases} 1 & \text{if } i\text{th draw is black} \\ 0 & \text{otherwise} \end{cases}$$

In the case where $N = 2$ and $m = 1$ what is $P(B_i) = 1$ for all i?

$$\Omega = \{BW, WB\}$$

$$P(B_1) = 1/2, \; P(B_2) = 1/2$$

$$P(W_1) = 1/2, \; P(W_2) = 1/2$$

Hypergeometric Random Variable - $E(X)$ - cont.

What about when $N = 3$ and $m = 1$?

$$\Omega = \{BW_1W_2, BW_2W_1, W_1BW_2, W_2BW_1, W_1W_2B, W_2W_1B\}$$

$$P(B_1) = 1/3, \; P(B_2) = 1/3, \; P(B_3) = 1/3$$

$$P(W_1) = 2/3, \; P(W_2) = 2/3, \; P(W_3) = 2/3$$

Proposition

$$P(B_i = 1) = m/N \text{ for all } i$$

Hypergeometric Random Variable - $E(X)$ - cont.

Let $X \sim \text{Hypergeo}(N, m, n)$ then $X = B_1 + B_2 + \cdots + B_n$

Hypergeometric Random Variable - $E(X)$ - 2nd way

Let $X \sim$ Hypergeo(N, m, n), what is $E(X)$?

Hypergeometric Random Variable - $Var(X)$

Let $X \sim$ Hypergeo(N, m, n), what is $Var(X)$?

$$\begin{aligned} Var(X) = Var\left(\sum_{i=1}^{n} B_i\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(B_i, B_j) \\ &= \sum_{i=1}^{n} Var(B_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} Cov(B_i, B_j) \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^{n} Var(B_i) &= \sum_{i=1}^{n} \frac{m}{N}\left(1 - \frac{m}{N}\right) \\ &= \sum_{i=1}^{n} \frac{m(N-m)}{N^2} = \frac{nm(N-m)}{N^2} \end{aligned}$$

Hypergeometric Random Variable - Variance, cont.

$$\begin{aligned} \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} Cov(B_i, B_j) &= \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} E(B_i B_J) - E(B_i)E(B_j) \\ &= \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} P(B_i = 1 \cap B_j = 1) - P(B_i = 1)P(B_j = 1) \\ &= \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} P(B_i = 1)P(B_j = 1 | B_i = 1) - \frac{m}{N}\frac{m}{N} \\ &= \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} \frac{m}{N}\frac{m-1}{N-1} - \frac{m}{N}\frac{m}{N} = \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} \frac{Nm(m-1) - m^2(N-1)}{N^2(N-1)} \\ &= \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} - \frac{m(N-m)}{N^2(N-1)} = -n(n-1)\frac{m(N-m)}{N^2(N-1)} \end{aligned}$$

Hypergeometric Random Variable - Variance, cont.

Let $X \sim$ Hypergeo(N, m, n)

$$\begin{aligned} Var(X) &= \sum_{i=1}^{n} Var(B_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} Cov(B_i, B_j) \\ &= \frac{nm(N-m)}{N^2} - n(n-1)\frac{m(N-m)}{N^2(N-1)} \\ &= \frac{nm(N-m)(N-1) - nm(n-1)(N-m)}{N^2(N-1)} \\ &= \frac{nm(N-m)}{N^2}\frac{N-n}{N-1} \\ &= n\frac{m}{N}\left(1 - \frac{m}{N}\right)\frac{N-n}{N-1} \end{aligned}$$

Poisson Random Variable - $E(X)$

Let $X \sim \text{Poisson}(\lambda)$, what is $E(X)$?

Poisson Random Variable - $Var(X)$

Let $X \sim \text{Poisson}(\lambda)$, what is $Var(X)$?

St. Petersburg Lottery

We start with \$1 on the table and a coin.

At each step: Toss the coin; if it shows Heads, take the money. If it shows Tails, I double the money on the table.

Let X be the amount you win, what is $E(X)$?

Moments

Some definitions,

Raw moment:

$$\mu'_n = E(X^n)$$

Central moment:

$$\mu_n = E[(X - \mu)^n]$$

Normalized / Standardized moment:

$$\frac{\mu_n}{\sigma^n}$$

Common Moments of Interest

Zeroth Moment:

$$\mu_0' = \mu_0 = 1$$

First Moment:

$$\mu_1' = E(X) = \mu$$
$$\mu_1 = E(X - \mu) = 0$$

Second Moment:

$$\mu_2 = E[(X-\mu)^2] = Var(X)$$
$$\mu_2' - (\mu_1')^2 = Var(X)$$

Third Moment:

$$\text{Skewness}(X) = \frac{\mu_3}{\sigma^3}$$

Fourth Moment:

$$\text{Kurtosis}(X) = \frac{\mu_4}{\sigma^4}$$
$$\text{Ex. Kurtosis}(X) = \frac{\mu_4}{\sigma^4} - 3$$

Note that some moments do not exist, which is the case when $E(X^n)$ does not converge.

Third and Forth Central Moments

$$\begin{aligned}\mu_3 &= E[(X-\mu)^3] = E(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3) \\ &= E(X^3) - 3\mu E(X^2) + 3\mu^3 - \mu^3 \\ &= E(X^3) - 3\mu\sigma^2 - \mu^3 \\ &= \mu_3' - 3\mu\sigma^2 - \mu^3\end{aligned}$$

$$\begin{aligned}\mu_4 &= E[(X-\mu)^4] = E(X^4 - 4X^3\mu + 6X^2\mu^2 - 4X\mu^3 + \mu^4) \\ &= E(X^4) - 4\mu E(X^3) + 6\mu^2 E(X^2) - 4\mu^4 + \mu^4 \\ &= E(X^4) - 4\mu E(X^3) + 2\mu^2(\sigma^2 + \mu^2) + 4\mu^2\sigma^2 + \mu^4 \\ &= \mu_4' - 4\mu\mu_3' + 6\mu^2\sigma^2 + 3\mu^4\end{aligned}$$

Moment Generating Function

The moment generating function of a discrete random variable X is defined for all real values of t by

$$M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} P(X = x)$$

This is called the moment generating function because we can obtain the moments of X by successively differentiating $M_X(t)$ wrt t and then evaluating at $t = 0$.

$$\begin{aligned}M_X(0) &= E[e^0] = 1 = \mu_0' \\ M_X'(t) &= \frac{d}{dt}E[e^{tX}] = E\left[\frac{d}{dt}e^{tX}\right] = E[Xe^{tX}] \\ M_X'(0) &= E[Xe^0] = E[X] = \mu_1' \\ M_X''(t) &= \frac{d}{dt}M_X'(t) = \frac{d}{dt}E[Xe^{tX}] = E\left[\frac{d}{dt}(Xe^{tX})\right] = E[X^2e^{tX}] \\ M_X''(0) &= E[X^2e^0] = E[X^2] = \mu_2'\end{aligned}$$

Moment Generating Function - Poisson

Let $X \sim \text{Pois}(\lambda)$ then

Moment Generating Function - Poisson Skewness

Moment Generating Function - Poisson Kurtosis

Moment Generating Function - Binomial

Let $X \sim \text{Binom}(n, p)$ then

$$
\begin{aligned}
M_X(t) = E[e^{tX}] &= \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \\
&= [pe^t + (1-p)]^n
\end{aligned}
$$

$$
\begin{aligned}
M_X'(t) &= npe^t(pe^t + 1 - p)^{n-1} \\
M_X'(0) &= \mu_1' = np = E(X)
\end{aligned}
$$

$$
\begin{aligned}
M_X''(t) &= n(n-1)(pe^t)^2(pe^t + 1 - p)^{n-2} + npe^t(pe^t + 1 - p)^{n-1} \\
M_X''(0) &= \mu_2' = n(n-1)p^2 + np = E(X^2)
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= \mu_2' - (\mu_1')^2 = n(n-1)p^2 + np - n^2p^2 \\
&= np((n-1)p + 1 - np) = np(1-p)
\end{aligned}
$$

Moment Generating Function - Normal

Let $Z \sim \mathcal{N}(0,1)$ and $X \sim \mathcal{N}(\mu, \sigma^2)$ where $X = \mu + \sigma Z$ then

$$
\begin{aligned}
M_Z(t) = E[e^{tZ}] &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2 - tx}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2} + \frac{t^2}{2}} dx \\
&= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} dx \\
&= e^{t^2/2}
\end{aligned}
$$

$$
\begin{aligned}
M_X(t) = E[e^{tX}] &= E[e^{t(\mu + \sigma Z)}] = E[e^{t\mu} e^{t\sigma Z}] \\
&= e^{t\mu} E[e^{t\sigma Z}] = e^{t\mu} M_U(t\sigma) \\
&= e^{t\mu} e^{t^2\sigma^2/2} = \exp\left(t\mu + \frac{t^2\sigma^2}{2}\right)
\end{aligned}
$$

Moment Generating Function - Normal, cont.

$$M'_X(t) = \exp(\mu t + \frac{t^2\sigma^2}{2})(\mu + t\sigma^2)$$
$$M'_X(0) = \mu'_1 = \mu$$

$$M''_X(t) = \exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)\sigma^2 + \exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)(\mu + t\sigma^2)^2$$
$$M''_X(0) = \mu'_2 = \mu^2 + \sigma^2$$

Moment Generating Function - Normal, cont.

$$\mu'_3 = M'''_X(0) = \mu^3 + 3\mu\sigma^2$$
$$\mu'_4 = M''''_X(0) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

$$\text{Skewness}(X) = \frac{\mu_3}{\sigma^3} = \frac{\mu'_3 - 3\mu\sigma^2 - \mu^3}{\sigma^3}$$
$$= \frac{\mu^3 + 3\mu\sigma^2 - 3\mu\sigma^2 - \mu^3}{\sigma^3} = 0$$

$$\text{Kurtosis}(X) = \frac{\mu_4}{\sigma^4} = \frac{\mu'_4 - 4\mu\mu'_3 + 6\mu^2\sigma^2 + 3\mu^4}{\sigma^4}$$
$$= \frac{\mu'_4 - 4\mu(\mu^3 + 3\mu\sigma^2) + 6\mu^2\sigma^2 + 3\mu^4}{\sigma^4}$$
$$= \frac{(\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4) - 6\mu^2\sigma^2 - 1\mu^4}{\sigma^4} = \frac{3\sigma^4}{\sigma^4} = 3$$

$$\text{Ex. Kurtosis}(X) = 0$$