Properties of Expected values and Variance

Christopher Croke

University of Pennsylvania

Math 115
UPenn, Fall 2011

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$.

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$. It turns out (and we have already used) that

$$E(r(X)) = \int_{-\infty}^{\infty} r(x)f(x)dx.$$

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$. It turns out (and we have already used) that

$$E(r(X)) = \int_{-\infty}^{\infty} r(x)f(x)dx.$$

This is not obvious since by definition $E(r(X)) = \int_{-\infty}^{\infty} xf_Y(x)dx$ where $f_Y(x)$ is the probability density function of $Y = r(X)$.

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$. It turns out (and we have already used) that

$$E(r(X)) = \int_{-\infty}^{\infty} r(x)f(x)dx.$$

This is not obvious since by definition $E(r(X)) = \int_{-\infty}^{\infty} xf_Y(x)dx$ where $f_Y(x)$ is the probability density function of $Y = r(X)$. You get from one integral to the other by careful uses of u substitution.

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$. It turns out (and we have already used) that

$$E(r(X)) = \int_{-\infty}^{\infty} r(x)f(x)dx.$$

This is not obvious since by definition $E(r(X)) = \int_{-\infty}^{\infty} xf_Y(x)dx$ where $f_Y(x)$ is the probability density function of $Y = r(X)$.
You get from one integral to the other by careful uses of u substitution.
One consequence is

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)f(x)dx = aE(X) + b.$$

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$. It turns out (and we have already used) that

$$E(r(X)) = \int_{-\infty}^{\infty} r(x)f(x)dx.$$

This is not obvious since by definition $E(r(X)) = \int_{-\infty}^{\infty} xf_Y(x)dx$ where $f_Y(x)$ is the probability density function of $Y = r(X)$.
You get from one integral to the other by careful uses of u substitution.
One consequence is

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)f(x)dx = aE(X) + b.$$

(It is not usually the case that $E(r(X)) = r(E(X))$.)

Expected value

Consider a random variable $Y = r(X)$ for some function r, e.g. $Y = X^2 + 3$ so in this case $r(x) = x^2 + 3$. It turns out (and we have already used) that

$$E(r(X)) = \int_{-\infty}^{\infty} r(x)f(x)dx.$$

This is not obvious since by definition $E(r(X)) = \int_{-\infty}^{\infty} xf_Y(x)dx$ where $f_Y(x)$ is the probability density function of $Y = r(X)$. You get from one integral to the other by careful uses of u substitution.

One consequence is

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)f(x)dx = aE(X) + b.$$

(It is not usually the case that $E(r(X)) = r(E(X))$.)

Similar facts old for discrete random variables.

Problem For X the uniform distribution on $[0, 2]$ what is $E(X^2)$?

Problem For X the uniform distribution on $[0,2]$ what is $E(X^2)$?

If $X_1, X_2, X_3, ...X_n$ are random variables and
$Y = r(X_1, X_2, X_3, ...X_n)$ then

$$E(Y) = \int\int\int ... \int r(x_1, x_2, x_3, ..., x_n) f(x_1, x_2, x_3, ..., x_n) dx_1 dx_2 dx_3 ... dx_n$$

where $f(x_1, x_2, x_3, ..., x_n)$ is the joint probability density function.

Problem For X the uniform distribution on $[0,2]$ what is $E(X^2)$?

If $X_1, X_2, X_3, ...X_n$ are random variables and $Y = r(X_1, X_2, X_3, ...X_n)$ then

$$E(Y) = \int\int\int ... \int r(x_1, x_2, x_3, ..., x_n) f(x_1, x_2, x_3, ..., x_n) dx_1 dx_2 dx_3 ... dx_n$$

where $f(x_1, x_2, x_3, ..., x_n)$ is the joint probability density function.
Problem Consider again our example of randomly choosing a point in $[0,1] \times [0,1]$.

Problem For X the uniform distribution on $[0,2]$ what is $E(X^2)$?

If $X_1, X_2, X_3, ...X_n$ are random variables and
$Y = r(X_1, X_2, X_3, ...X_n)$ then

$$E(Y) = \int\int\int ... \int r(x_1, x_2, x_3, ..., x_n) f(x_1, x_2, x_3, ..., x_n) dx_1 dx_2 dx_3 ... dx_n$$

where $f(x_1, x_2, x_3, ..., x_n)$ is the joint probability density function.
Problem Consider again our example of randomly choosing a point in $[0,1] \times [0,1]$. We could let X be the random variable of choosing the first coordinate and Y the second. What is $E(X+Y)$?
(note that $f(x,y) = 1$.)

Problem For X the uniform distribution on $[0,2]$ what is $E(X^2)$?

If $X_1, X_2, X_3, ...X_n$ are random variables and
$Y = r(X_1, X_2, X_3, ...X_n)$ then

$$E(Y) = \int\int\int ... \int r(x_1, x_2, x_3, ..., x_n) f(x_1, x_2, x_3, ..., x_n) dx_1 dx_2 dx_3 ... dx_n$$

where $f(x_1, x_2, x_3, ..., x_n)$ is the joint probability density function.
Problem Consider again our example of randomly choosing a point in $[0,1] \times [0,1]$. We could let X be the random variable of choosing the first coordinate and Y the second. What is $E(X+Y)$?
(note that $f(x,y) = 1$.)

Easy properties of expected values:
If $Pr(X \geq a) = 1$ then $E(X) \geq a$.
If $Pr(X \leq b) = 1$ then $E(X) \leq b$.

Properties of $E(X)$

A little more surprising (but not hard and we have already used):

$$E(X_1 + X_2 + X_3 + \dots + X_n) = E(X_1) + E(X_2) + E(X_3) + \dots + E(X_n).$$

Properties of $E(X)$

A little more surprising (but not hard and we have already used):

$$E(X_1 + X_2 + X_3 + ... + X_n) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n).$$

Another way to look at binomial random variables;
Let X_i be 1 if the i^{th} trial is a success and 0 if a failure.

A little more surprising (but not hard and we have already used):

$$E(X_1 + X_2 + X_3 + ... + X_n) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n).$$

Another way to look at binomial random variables;
Let X_i be 1 if the i^{th} trial is a success and 0 if a failure. Note that $E(X_i) = 0 \cdot q + 1 \cdot p = p$.

Properties of $E(X)$

A little more surprising (but not hard and we have already used):

$E(X_1 + X_2 + X_3 + ... + X_n) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n)$.

Another way to look at binomial random variables;
Let X_i be 1 if the i^{th} trial is a success and 0 if a failure. Note that $E(X_i) = 0 \cdot q + 1 \cdot p = p$.
Our binomial variable (the number of successes) is $X = X_1 + X_2 + X_3 + ... + X_n$ so

$$E(X) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n) = np.$$

Properties of $E(X)$

A little more surprising (but not hard and we have already used):

$$E(X_1 + X_2 + X_3 + ... + X_n) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n).$$

Another way to look at binomial random variables;
Let X_i be 1 if the i^{th} trial is a success and 0 if a failure. Note that $E(X_i) = 0 \cdot q + 1 \cdot p = p$.
Our binomial variable (the number of successes) is
$X = X_1 + X_2 + X_3 + ... + X_n$ so

$$E(X) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n) = np.$$

What about products?

Properties of $E(X)$

A little more surprising (but not hard and we have already used):

$E(X_1 + X_2 + X_3 + ... + X_n) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n)$.

Another way to look at binomial random variables;
Let X_i be 1 if the i^{th} trial is a success and 0 if a failure. Note that $E(X_i) = 0 \cdot q + 1 \cdot p = p$.
Our binomial variable (the number of successes) is $X = X_1 + X_2 + X_3 + ... + X_n$ so

$$E(X) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_n) = np.$$

What about products? Only works out well if the random variables are **independent**. If $X_1, X_2, X_3, ...X_n$ are independent random variables then:

$$E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i).$$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3){=}0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information!

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3){=}0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)
- $\sigma^2(aX + b) = a^2\sigma^2(X)$.

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)
- $\sigma^2(aX + b) = a^2\sigma^2(X)$. Proof: $\sigma^2(aX + b) =$

  $E[(aX+b-(a\mu+b))^2]$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)
- $\sigma^2(aX + b) = a^2\sigma^2(X)$. Proof: $\sigma^2(aX + b) =$

  $$E[(aX+b-(a\mu+b))^2] = E[(aX-a\mu)^2]$$

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)
- $\sigma^2(aX + b) = a^2\sigma^2(X)$. Proof: $\sigma^2(aX + b) =$

  $$E[(aX+b-(a\mu+b))^2] = E[(aX-a\mu)^2] = a^2E[(X-\mu)^2]$$

Properties of $Var(X)$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3)=0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)
- $\sigma^2(aX + b) = a^2\sigma^2(X)$. Proof: $\sigma^2(aX + b) =$

  $$E[(aX+b-(a\mu+b))^2] = E[(aX-a\mu)^2] = a^2E[(X-\mu)^2] = a^2\sigma^2(X).$$

Problem: Consider **independent** random variables X_1, X_2, and X_3, where $E(X_1) = 2$, $E(X_2) = -1$, and $E(X_3){=}0$. Compute $E(X_1^2(X_2 + 3X_3)^2)$.
There is not enough information! Also assume $E(X_1^2) = 3$, $E(X_2^2) = 1$, and $E(X_3^2) = 2$.

Facts about $Var(X)$:

- $Var(X) = 0$ means the same as: there is a c such that $Pr(X = c) = 1$.
- $\sigma^2(X) = E(X^2) - E(X)^2$ (alternative definition)
- $\sigma^2(aX + b) = a^2\sigma^2(X)$. Proof: $\sigma^2(aX + b) =$

  $$E[(aX+b-(a\mu+b))^2] = E[(aX-a\mu)^2] = a^2E[(X-\mu)^2] = a^2\sigma^2(X).$$

- For **independent** $X_1, X_2, X_3, ..., X_n$

  $$\sigma^2(X_1+X_2+X_3+...+X_n) = \sigma^2(X_1)+\sigma^2(X_2)+\sigma^2(X_3)+...+\sigma^2(X_n).$$

Properties of $Var(X)$

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + ... + a_n^2\sigma^2(X_n).$$

Properties of $Var(X)$

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + ... + a_n^2\sigma^2(X_n).$$

Now we can compute the variance of the binomial distribution with parameters n and p.

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + ... + a_n^2\sigma^2(X_n).$$

Now we can compute the variance of the binomial distribution with parameters n and p. As before $X = X_1 + X_2 + ... + X_n$ where the X_i are independent with $Pr(X_i = 0) = q$ and $Pr(X_i = 1) = p$.

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + ... + a_n^2\sigma^2(X_n).$$

Now we can compute the variance of the binomial distribution with parameters n and p. As before $X = X_1 + X_2 + ... + X_n$ where the X_i are independent with $Pr(X_i = 0) = q$ and $Pr(X_i = 1) = p$.
So $\mu(X_i) = p$

Properties of $Var(X)$

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + ... + a_n^2\sigma^2(X_n).$$

Now we can compute the variance of the binomial distribution with parameters n and p. As before $X = X_1 + X_2 + ... + X_n$ where the X_i are independent with $Pr(X_i = 0) = q$ and $Pr(X_i = 1) = p$.
So $\mu(X_i) = p$ and
$\sigma^2(X_i) = E(X_i^2) - \mu(X_i)^2 =$

Properties of $Var(X)$

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + ... + a_n^2\sigma^2(X_n).$$

Now we can compute the variance of the binomial distribution with parameters n and p. As before $X = X_1 + X_2 + ... + X_n$ where the X_i are independent with $Pr(X_i = 0) = q$ and $Pr(X_i = 1) = p$.
So $\mu(X_i) = p$ and
$\sigma^2(X_i) = E(X_i^2) - \mu(X_i)^2 = p - p^2 = p(1-p) = pq$.

Properties of $Var(X)$

Note that the last statement tells us that

$$\sigma^2(a_1X_1 + a_2X_2 + a_3X_3 + \ldots + a_nX_n) =$$

$$= a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + a_3^2\sigma^2(X_3) + \ldots + a_n^2\sigma^2(X_n).$$

Now we can compute the variance of the binomial distribution with parameters n and p. As before $X = X_1 + X_2 + \ldots + X_n$ where the X_i are independent with $Pr(X_i = 0) = q$ and $Pr(X_i = 1) = p$.
So $\mu(X_i) = p$ and
$\sigma^2(X_i) = E(X_i^2) - \mu(X_i)^2 = p - p^2 = p(1-p) = pq$.
Thus $\sigma^2(X) = \Sigma\sigma^2(X_i) = npq$.

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0, 1] \times [0, 1]$.

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0,1] \times [0,1]$. $Z = X + Y$ where X and Y both represent choosing a point randomly from $[0,1]$.

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0,1] \times [0,1]$. $Z = X + Y$ where X and Y both represent choosing a point randomly from $[0,1]$. They are independent and last time we showed $\sigma^2(X) = \frac{1}{12}$.

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0,1] \times [0,1]$. $Z = X + Y$ where X and Y both represent choosing a point randomly from $[0,1]$. They are independent and last time we showed $\sigma^2(X) = \frac{1}{12}$. So $\sigma^2(Z) = \frac{1}{6}$.

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0,1] \times [0,1]$. $Z = X + Y$ where X and Y both represent choosing a point randomly from $[0,1]$. They are independent and last time we showed $\sigma^2(X) = \frac{1}{12}$. So $\sigma^2(Z) = \frac{1}{6}$.

What is $E(XY)$?

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0,1] \times [0,1]$. $Z = X + Y$ where X and Y both represent choosing a point randomly from $[0,1]$. They are independent and last time we showed $\sigma^2(X) = \frac{1}{12}$. So $\sigma^2(Z) = \frac{1}{6}$.

What is $E(XY)$? They are independent so $E(XY) = E(X)E(Y) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Properties of $Var(X)$

Consider our random variable Z which is the sum of the coordinates of a point randomly chosen from $[0,1] \times [0,1]$. $Z = X + Y$ where X and Y both represent choosing a point randomly from $[0,1]$. They are independent and last time we showed $\sigma^2(X) = \frac{1}{12}$. So $\sigma^2(Z) = \frac{1}{6}$.

What is $E(XY)$? They are independent so
$E(XY) = E(X)E(Y) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.
($\sigma^2(XY)$ is more complicated.)

Why is $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$ even when not independent?

Why is $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$ even when not independent?

$$E(\sum_{i=1}^{n} X_i) = \int \int ... \int (x_1 + x_2 + ... + x_n) f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n$$

Why is $E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i)$ even when not independent?

$$E(\sum_{i=1}^n X_i) = \int\int ... \int (x_1 + x_2 + ... + x_n) f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n$$

$$= \sum_{i=1}^n \int\int ... \int x_i f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n.$$

Why is $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$ even when not independent?

$$E(\sum_{i=1}^{n} X_i) = \int \int ... \int (x_1 + x_2 + ... + x_n) f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n$$

$$= \sum_{i=1}^{n} \int \int ... \int x_i f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n.$$

$$= \sum_{i=1}^{n} \int x_i f_i(x_i) dx_i = \sum_{i=1}^{n} E(X_i).$$