Expected Value of Functions of a Random Variable

C. Durso

Expectation of Functions

Definition

If X is a discrete random variable $(S \subseteq \mathbb{R}, M, P)$ where P is induced by a density f, and $g: S \to \mathbb{R}$, the expected value of $g(X)$ equals $\sum_{x \in S} g(x) f(x)$. It is denoted $E[g(X)]$.

Definition

If X is a continuous random variable $(S \subseteq \mathbb{R}, M, P)$ where P is induced by a density f, then if $g: S \to \mathbb{R}$ is a continuous function, the expected value of $g(X)$ equals $\int g(x) f(x)\, dx$. It is denoted $E[g(X)]$.

(In fact, these definitions coincide with the mean of the random variable $Y = g(X)$. This is relatively easy to see in the discrete case, but is true generally.)

Example $E_X(g(X)) = E_Y(Y)$

- Let (S, M, P) be the model for rolling a fair die and recording the number rolled.
- Let g be the random variable defined by $g(1) = 0$, $g(2) = g(3) = 1$, $g(4) = g(5) = 2$, and $g(6) = 3$
- Then $E_X[g(X)] = \sum_{k=1}^{6} g(k)\frac{1}{6} = 0\left(\frac{1}{6}\right) + 1\left(\frac{1}{6} + \frac{1}{6}\right) + 2\left(\frac{1}{6} + \frac{1}{6}\right) + 3\left(\frac{1}{6}\right) = \frac{3}{2}$

Example $E_X(g(X)) = E_Y(Y)$

Now consider the probability space Y induced by g:

- The sample space is $\{0,1,2,3\}$.
- $f(0) = \frac{1}{6}$, $f(1) = \frac{1}{6} + \frac{1}{6}$, $f(1) = \frac{1}{6} + \frac{1}{6}$,
  $f(2) = \frac{1}{6} + \frac{1}{6}$, and $f(3) = \frac{1}{6}$
- Then $E_Y[Y] = \sum_{j=0}^{3} j f(j) = 0\left(\frac{1}{6}\right) + 1\left(\frac{1}{6} + \frac{1}{6}\right) + 2\left(\frac{1}{6} + \frac{1}{6}\right) + 3\left(\frac{1}{6}\right) = \frac{3}{2}$

Basics of Expectation

Theorem

If X is a random variable, c is a constant, then

- $E[cX] = cE[X]$
- $E[g(X) + h(X)] = E[g(X)] + E[h(X)]$

Note: $E[X + c] = E[X] + c$

Mean Using Expectation

Fact

If X is a random variable, $\bar{X} = E[X]$.

Variance Using Expectation

Fact

If X is a random variable,
$Var[X] = E[(X - \bar{X})^2] = \mathrm{E}[X^2] - (E[X])^2$.

Jointly Distributed Random Variables

C. Durso

Jointly Distributed Random Variables

Definition

If X_1, X_2,...X_n are random variables on a common probability space, (S, M, P) in the sense that $X_i: S \rightarrow \mathbb{R}$ for $i \in \{1,2, \dots n\}$, then X_1, X_2,...X_n are *jointly distributed* random variables.

Jointly Distributed Projection Variables

Fact

If the sample space S of a probability space (S, M, P) is a subset of $\mathbb{R}^n$, then the projection functions $X_i, i \in \{1,2, \dots n\}$ defined by $X_i\big((x_1, x_2, \dots x_n)\big) = x_i$ are jointly distributed random variables.

Each random variable gives rise to a probability space in the usual way, called its marginal distribution.

Probability Space From Jointly Distributed Random Variables

Example

Random variables $X_1, X_2,...X_n$ on a common probability space, (S, M, P), generate a new probability space (S', M', P'):

- The set $S' = \{(x_1, x_2, \ldots x_n) \in \mathbb{R}^n | \exists s \in S\ with\ (x_1, x_2, \ldots x_n) = (X_1(s), X_2(s), \ldots X_n(s))\}$
- The set M' a valid collection of measurable sets including all rectangles $[a_1, b_1]\times[a_2, b_2]\times \cdots [a_n, b_n]$
- The probability function P' defined by the extension from $P'([a_1, b_1]\times[a_2, b_2]\times \cdots [a_n, b_n]) = P(\{s|(X_1(s), X_2(s), \ldots X_n(s)) \in [a_1, b_1]\times[a_2, b_2]\times \cdots [a_n, b_n]\})$

Probability Space From Jointly Distributed Random Variables

Property

(S', M', P') is a probability space with the projection distributions distributed as the original $X_1, X_2, \ldots X_n$.

Example of Construction of Jointly Distributed Random Variables

- Consider a probability space defined by outcomes “$disagree$,” “$somewhat\ disagree$,” “$somewhat\ agree$,” and “$agree$” with probabilities $0.1, 0.2, 0.4$, and 0.3 respectively.
- Define X by $X(agree) = X(disagree) = 1$, and otherwise $X(s) = 0$.
- Define Y by $Y(\ disagree) = 0$, $Y(somewhat\ disagree)\ 1$, $Y(somewhat\ agree) = 2$, and $Y(agree) = 3$.

Continuous Probability Space in $\mathbb{R}^n$

Definition (sketch)

If the sample space S of a probability space (S, M, P) is a subset of $\mathbb{R}^n$ and there exists a measurable function $f: \mathbb{R}^n \to [0, \infty)$ such that for any Cartesian product $A = (a_1, b_1) \times (a_2, b_2) \times \cdots (a_n, b_n)$,

$$P(A) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_n}^{b_n} f(x_2, x_2, \dots x_n) dx_n \dots dx_2 dx_1$$

then (S, M, P) is a continuous probability space.

Independent Random Variables

Definition

Random variables X and Y on a probability space $(S \subseteq \mathbb{R}, M, P)$ are independent if, given any A of the form $X^{-1}(A')$ for A' an interval and B of the form $Y^{-1}(B')$ for B' an interval, the events $\{(x, y) | x \in A\}$ and $\{(x, y) | y \in B\}$ are independent.

Mutually Independent Random Variables

Definition

Random variables $X_1, X_2, \ldots X_n$ on a probability space $(S \subseteq \mathbb{R}, M, P)$ are mutually independent if, given any A_{i} of the form $X^{-1}(A_i')$ for A_i' an interval, $i \in \{1, \ldots n\}$, the events $\{x_i | x \in A_i\}$ are mutually independent.

Independent Random Variables Discrete Example

Example

Consider the discrete probability space that models rolling a fair die twice independently and recording the results in order. The sample space is $S = \{1,2,3,4,5,6\}\times\{1,2,3,4,5,6\}$. The collection of measurable sets M is the power set of S. The density function f is defined by $f\big((a,b)\big) = \frac{1}{36}$.

- Then the functions $X: (x,y) \to x$ and $Y: (x,y) \to y$ are jointly distributed random variables.
- They are independent. Given $A \subseteq \{1,2,3,4,5,6\}$ and $B \subseteq \{1,2,3,4,5,6\}$, the probability of the event $\{(x,y)|x \in A\}$ equals $\frac{|A|}{6}$, the probability of the event $\{(x,y)|y \in B\}$ equals $\frac{|B|}{6}$, and the probability of the intersection of $\{(x,y)|x \in A\}$ and $\{(x,y)|y \in B\}$ equals $\frac{|A||B|}{36}$.

Discrete Example Not Independent

Example

Consider a discrete probability space that models rolling a die twice, with a higher probability of repeating values, and recording the results in order. The sample space is $S = \{1,2,3,4,5,6\}\times\{1,2,3,4,5,6\}$. The collection of measurable sets M is the power set of S. Define the density function f by $f\big((a,b)\big) = \begin{cases} \frac{2}{42} & if\ a=b \\ \frac{1}{42} & if\ a\neq b \end{cases}$.

- Then the functions $X: (x,y) \to x$ and $Y: (x,y) \to y$ are jointly distributed random variables.
- They are not independent. Let A be the event $X^{-1}(\{6\})$ and let B be the event $Y^{-1}(\{1\})$. Each event has probability $\frac{2}{42} + \frac{5}{42} = \frac{1}{6}$. The probability of the intersection, $\{6,1\}$, is $\frac{1}{42} \neq \frac{1}{36}$.

General Example of Independent Random Variables: Discrete Version

Fact

If $X = (S_X \subseteq \mathbb{R}, M_X, P_X)$ and $Y = (S_Y \subseteq \mathbb{R}, M_Y, P_Y)$ are discrete random variables with densities and f_X and f_Y, then there exists a discrete probability space $(S_X \times S_Y, M, P)$ in which P is induced by

$f\left((x_i, y_j)\right) = f_X(x_i) f_Y(y_j)$.

General Example of Independent Random Variables: Continuous Version

Fact

If $X = (S_X \subseteq \mathbb{R}, M_X, P_X)$ and $Y = (S_Y \subseteq \mathbb{R}, M_Y, P_Y)$ are independent random variables with densities and f_X and f_Y, then there exists a probability space $(S_X \times S_Y, M, P\ )$ in which P is induced by $P(\{(x, y) | x \in A \wedge y \in B\}) = \int_{y \in B} \int_{x \in A} f_X(x) f_Y(y) dx dy$

Quick check: $\int_{y \in B} \int_{x \in A} f_X(x) f_Y(y) dx dy = \int_{y \in B} f_Y(y) \int_{x \in A} f_X(x) dx dy = \int_{y \in B} f_Y(y) P_X(A) dy = P_X(A) P_Y(B)$, as required by independence.

Some Joint Expectations

C. Durso

Expected Value of Sum Discrete Version

Let $X = (S_X \subseteq \mathbb{R}, M_X, P_X)$ and $Y = (S_Y \subseteq \mathbb{R}, M_Y, P_Y)$ be jointly distributed discrete random variables with marginal densities and f_X and f_Y. Let (S, M, P) be the underlying discrete probability space with density function f. Denote expected value in this space by E.

Theorem

If X and Y are jointly distributed discrete random variables, $E[X + Y] = E_X[X] + E_Y[Y]$.

$E[X+Y] = E_X[X] + E_Y[Y]$, Discrete Case

$$E[X+Y] = \sum_{s \in S} \big(X(s) + Y(s)\big) f(s)$$

$$= \sum_{s \in S} X(s) f(s) + \sum_{s \in S} Y(s) f(s)$$

$$= \sum_{x \in S_X} \sum_{s \in S, X(s)=x} x f(s) + \sum_{y \in S_Y} \sum_{s \in S, Y(s)=y} y f(s)$$

$$= \sum_{x \in S_X} x \sum_{s \in S, X(s)=x} f(s) + \sum_{y \in S_Y} y \sum_{s \in S, Y(s)=y} f(s)$$

$$= \sum_{x \in S_X} x f_X(x) + \sum_{y \in S_Y} y f_Y(y)$$

$$= E_X[X] + E_Y[Y]$$

Expected Value of Sum
Continuous Version

Let $(\mathbb{R}^2, M, P)$ be a continuous probability space with density function f and let $X = (\mathbb{R}, M_X, P_X)$ and $Y = (\mathbb{R}, M_Y, P_Y)$ be the projection function onto the first and second components with densities and f_X and f_Y.

Theorem

If X and Y are jointly distributed continuous random variables, $E[X + Y] = E_X[X] + E_Y[Y]$.

Joint Expectation Identities Under Independence Continuous Version

Let $X = (\mathbb{R}, M_X, P_X)$ and $Y = (\mathbb{R}, M_Y, P_Y)$ be independent random variables with densities and f_X and f_Y, and let $(\mathbb{R}^2, M, P\ )$ be the probability space in which P is induced by $P(\{(x,y)|x \in A \wedge y \in B\}) = \int_{y\in B}\int_{x\in A} f_X(x) f_Y(y) dx dy$. Denote expectation in these spaces by E_X, E_Y, and $E_{X\times Y}$ respectively.

Theorem

The following identities hold for expected values in $X \times Y$:

- $E_{X\times Y}[XY] = E_X[X]E_Y[Y]$
- $Var_{X\times Y}[X+Y] = Var_X[X] + Var_Y[Y]$

Joint Expectation Identities Under Independence Discrete Version

Let $X = (\mathrm{S}_{\mathrm{X}}, M_X, P_X)$ and $Y = (\mathrm{S}_{\mathrm{Y}}, M_Y, P_Y)$ be independent random variables with densities and f_X and f_Y, and let $(S_X \times S_Y, M, P\ )$ be the probability space in which P is induced by the density function f with $f\big((x, y)\big) = f_X(x) f_Y(y)$. Denote expectation in these spaces by E_X, E_Y, and $E_{X\times Y}$ respectively.

Theorem

The following identities hold for expected values in $X\times Y$:

- $E_{X\times Y}[XY] = E_X[X] + E_Y[Y]$
- $Var_{X\times Y}[X + Y] = Var_X[X] + Var_Y[Y]$

$E_{X\times Y}[XY] = E_X[X]E_Y[Y]$, Continuous Case

$$\int\int xyf_X(x)f_Y(y)dxdy = \int yf_Y(y)\left(\int xf_X(x)dx\right)dy$$

$$= \int yf_Y(y)(E_X[X])dy$$

$$= E_X[X]\int yf_Y(y)dy = E_X[X]E_Y[Y]$$

$Var_{X\times Y}[X+Y] = Var_X[X] + Var_Y[Y]$

$Var_{X\times Y}[X+Y] = E_{X\times Y}[(X+Y)^2] - (E_{X\times Y}[X+Y])^2$

$= E_{X\times Y}[X^2 + 2XY + Y^2] - (E_X[X] + E_Y[Y])^2$

$= E_X[X^2] + 2E_X[X]E_Y[Y] + E_Y[Y^2] - E_X[X]^2 - 2E_X[X]E_Y[Y] - E_Y[Y]^2$

$= E_X[X^2] - E_X[X]^2 + E_Y[Y^2] - E_Y[Y]^2 + 2E_X[X]E_Y[Y] - 2E_X[X]E_Y[Y]$

$= Var_X[X] + Var_Y[Y]$

Binomial Mean and Variance

Example

C. Durso

Expected Value of $Binomial(n, p)$

Formula
$E[binomial(n,p)] = np$

Why?

Random variable $X \sim binomial(n, p)$

- Sum of n independent $binomial(1, p)$ random variables X_i
- $E[X_i] = 0(1-p) + 1(p) = p$
- $E[X] = \sum_{i=1}^{n} E[X_i] = np$

Variance of $Binomial(n, p)$

Formula
$Var[binomial(n, p)] = np(1 - p)$

Why?

Random variable $X \sim binomial(n, p)$

- Sum of n independent $binomial(1, p)$ random variables X_i
- $Var[X_i] = E\left[X_i^2\right] - E[X_i]^2 = 0^2(1-p) + 1^2(p) - \mathrm{p}^2 = \mathrm{p}(1-\mathrm{p})$
- $Var[X] = \sum_{i=1}^{n} Var[X_i] = np(1-p)$

Normal Approximation to Binomial

Example

C. Durso

Expected Value of Binomial(n,p)

Formula

$$E[binomial(n,p)] = np$$

Variance of Binomial(n,p)

Formula

$$Var[binomial(n,p)] = np(1-p)$$

Normal Approximation

- If $np(1-p) > 3$, say, $X \sim binomial(n, p)$, and $Y \sim normal\big(np, np(1-p)\big)$, then $P(X = k)$ is approximately $P(k - \frac{1}{2} \leq Y \leq k + \frac{1}{2})$.
- Note X and Y have the same mean and variance.

Mean and Variance of Sample Mean

C. Durso

Distribution of Mean

Theorem

Let $X_1, X_2, \dots X_n$ be independent random variables with the same distribution (iid), with mean μ and variance σ^2. Then $E\left[\frac{1}{n}\sum X_i\right] = \mu$ and $Var\left[\frac{1}{n}\sum X_i\right] = \frac{\sigma^2}{n}$.

$$Var\left[\sum X_i\right] = \sum Var(X_i) = n\sigma^2$$

$Var[cY] = E[(cY)^2] - (E[cY])^2 = c^2 Var[Y]$, so

$$Var\left[\frac{1}{n}\sum X_i\right] = \frac{1}{n^2} Var\left[\sum X_i\right]$$

$$= \frac{\sigma^2}{n}$$

Expected Value of Sample Variance

Theorem

Let $X_1, X_2, \ldots X_n$ be independent random variables with the same distribution (iid), with mean μ and variance σ^2. Then $E\left[\frac{1}{n-1}\sum\left(X_i - \frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] = \sigma^2$.

Overview of One Sample Test

C. Durso

Tests of Mean or Median

- $x_1, x_2, \dots x_n$ is a sample from a random variable X with median=μ
  - (sign test: $X_i = X_j$ not required, only need information on truth values of $x_i - \mu_0 > 0$ and $x_i - \mu_0 < 0$)
- Null hypothesis: $\mu = \mu_0$

Covered Tests

- z-test
- Student’s t-test
- Wilcoxon signed rank test
- Sign test

In order of most restrictive hypotheses on X to least

Some Applications

- Does sampled population have same mean as reference population?
  - Yield of new manufacturing method compared to old, well-studied method
  - Learning measure for students in new program compared to old, well-studied program
- Do pre- and post-intervention measurements differ systematically in size?
  - Compare medical condition pre- and post-treatment
  - Compare outcomes in matched pairs with one treated, one untreated

Which to Use?

- Apply the test with the most restrictive hypotheses that data meet.
- This test will be the most sensitive to violations of the null hypothesis.