270: Random Variables

D. Alex Hughes

Math Camp

September 21, 2012

To this point, we have avoided discussion of Random Variables. No longer.

Definition

A *random variable* is a function whose domain is a sample space and whose range is some set of real numbers.

If a random variable is denoted X and has as its domain the sample space $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, then $X(\omega_j)$ is the value of X at element ω_j.

That is, $X(\omega_j)$ is the number that the function rule (X) assigns to element $\omega_j \in \Omega$.

Random variables are exceedingly useful to store the values of stochastic processes. Most of the time, their use is intuitive, though this is not universally the case.

Vote Buying Example

Example

There are 100 senators. Each of these senators could potentially be involved in a vote buying scandal. There are 2^{100} possible ways the senate could/not be corrupt – {not, not, not, corrupt, corrupt, not, ... }

But we don't really care *who* – we're not Fox News or MSNBC – we just want the data.

So we take the values and store them, *in a random variable*. In this case, we let the RV be 1 if a senator was corrupt, and 0 otherwise. This way, $\Omega = [0, 100]$, not $[0, 2^{100}]$.

Preliminaries

Think back to our initial example of a discrete distribution: flipping a coin three times.

- Recall, what was the:
  - Experiment
  - Sample Outcome
  - Sample Space
  - Event
- If we repeated this experiment many times, we might want a convenient placeholder for our results.
- Create a random variable, denoted X, that takes the value 1 for each experiment in which the event occurs and takes the value 0 for each experiment in which the event doesn't occur.
- Then $\Omega = [0, n]$, not $[0, 2^n]$
- Quick toss to R $\rightarrow$

Distributional Statistics

There are several measurements we're interested with distributions:

1. Central Tendency:
   - Mean: $\left(\frac{1}{N}\right)\sum_{i=1}^{N} x_i$
   - Median: Select $\left(\frac{N}{2}\right)$ from Ordered Set
   - Mode: Maximally occurrent observation. Useful for nominal-level variables.
2. Dispersion:
   - Ordinal: Comparison of Median and Mode
   - Interval and Ratio: Variance
     - **Variance**: $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$
     - **Standard Deviation**: $\sigma = (\sigma^2)^{\frac{1}{2}}$
3. Skewness: $\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^3$
4. Kurtosis: $\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^4$

Expectation of Discrete Random Variables

Definition

Let X be a numerically-valued (i.e. not True/False, Blue/Green) random variable across Ω with distribution function $f(x)$.

Then, the *expected value*, denoted E[X] or μ, is defined as

$$E[X] = \sum_{x \in \Omega} x \cdot f(x).$$

Here, x refers to the value of the outcome, and $f(x)$ the probability function that maps that outcome. In conversation we refer to the expectation as the *mean*, and can be thought of like the center of gravity of the distribution.

Together with a measure of *dispersion* (and perhaps the other higher-order moments), the mean is a core description we are interested in for any variable.

Expectation of Discrete Random Varialbles

Example

Consider the game of roulette. After bets are placed, the croupier spins, the marble lands, and one of 38 (00, 0, 1, ... 36) to be the winner. Any of the numbers is equally likely. 00 and 0 are neither odd nor even. We bet \$1 on "odd" at even money. What is our expected earnings for each bet?

Answer

$$
\begin{aligned}
p_X(1) &= P(X=1) = \frac{18}{38} = \frac{9}{19} \\
p_X(-1) &= P(X=-1) = \frac{20}{38} = \frac{10}{19} \\
\text{“expected winnings”} &= \$1\left(\frac{9}{19}\right) - \$1\left(\frac{10}{19}\right) \\
&= -\frac{\mathbf{\$1}}{\mathbf{19}}
\end{aligned}
$$

Expectation of Discrete Random Variables

Example

You roll 12 times a fair 3-sided die (Dragonslayer?) that has the values 1, 2, and 4 on the sides. After each roll, you store the number showing in a random variable Q. If you had to place a bet, on what number would you bet is the sum of the rolls?

Answer

Our expectation for the sum is going to n*E[Q].

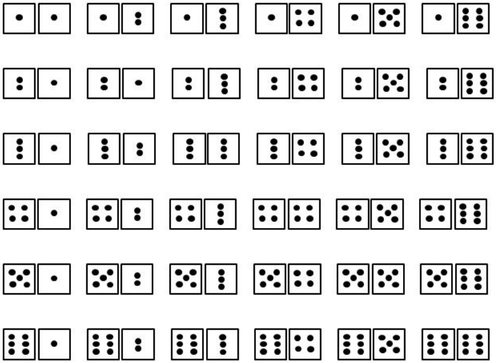
$$\begin{aligned} E[Q] &= \sum_{i=1}^{10} x \cdot f(x) \\ &= (1)(1/3) + 2(1/3) + (4)(1/3) \\ &= 7/3 \end{aligned}$$

And so the number most likely to come up is $12 * 7/3 = \mathbf{28}$.

Why does the house want 7s in craps?

Example

You throw two dice and add the values. What is the expected value? Why does the house want 7s?



Answer

Answer

The house wants to play the best odds – it wants to make you lose with the highest frequency.

$$\begin{aligned} E[X] &= \sum x \cdot f(x) \\ &= (1/36)(2) + (2/36)(3) + (3/36)(4) + (4/36)(5) + (5/36)(6) \\ &+ (6/36)(7) + (5/36)(8) + (4/36)(9) + (3/36)(10) + (2/36)(11) \\ &+ (1/36)(12) = 252/36 = \mathbf{7} \end{aligned}$$

Example

Suppose X is a binomial random variable with $p = \frac{5}{9}$ and $n = 3$. What is the expected value of X?

Binomial RV Answer

Answer

$$
\begin{aligned}
p_X(k) &= \binom{3}{k}\left(\frac{5}{9}\right)^k\left(\frac{4}{9}\right)^{3-k}, k = 0, 1, 2, 3 \\
E[X] &= \sum_{k=0}^{3} k \cdot \binom{3}{k}\left(\frac{5}{9}\right)^k\left(\frac{4}{9}\right)^{3-k} \\
&= (0)\left(\frac{64}{729}\right) + (1)\left(\frac{240}{729}\right) + \left(\frac{300}{729}\right) + (3)\left(\frac{125}{729}\right) \\
&= \frac{1215}{729} \\
&= 3\left(\frac{5}{9}\right) \text{(remember this result, we will come back to this later)}
\end{aligned}
$$

Make that Chedda'

Example

A fair coin is flipped until the first tail appears. You win $\$2^k$ where k is the toss on which the first tail appears. How much are you willing to pay to play this game?

Answer

This is the St. Petersburg paradox.

$$p_X(2^k) = P(X = 2^k) = \frac{1}{2^k}, k = 1, 2, ...$$

And so,

$$E[X] = \sum_{\forall k} 2^k p_X(2^k) = \sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k} = 1 + 1 + \cdots$$

Expectation of Continuous Random Variables

Definition

If Y is a continuous random variable with pdf $f_Y(y)$,

$$E[Y] = \int_{-\infty}^{\infty} y \cdot f_Y(y) dy$$

In the same was as the discrete case, one may think of this as a center of gravity of the distribution.

Example

A continuous uniform density function, f(x) = 1 is defined on the range [0,1]. What is the expected value of this function?

Answer

$$
\begin{aligned}
E[X] &= \int_0^1 x \cdot f(x)dx \\
&= \int_0^1 1x dx \\
&= \frac{1}{2}x^2|_0^1 \\
&= \frac{1}{2}(1-0) \\
&= \mathbf{0.5}
\end{aligned}
$$

Example

Suppose a human rights are violated in a way that is well described by the following pdf,

$$f_Y(y) = \frac{1}{\mu} e^{\frac{-y}{\mu}}, y \geq 0$$

where μ is a positive constant known as the mean global violation rate. What is the expected number of violations?

Answer

$$E[Y] = \int_0^\infty y \frac{1}{\mu} e^{\frac{-y}{\mu}} dy$$

Let $w = y/\mu$ and $dw = 1/\mu dy$. Then, $E[Y] = \mu \int_0^\infty we^{-w} dw$. Let $u = w$ and $dv = e^{-w} dw$, then integrating by parts gives,

$$E[Y] = \mu[-we^{-w} - e^{-w}]_0^\infty = \mu$$

Expectation Identities

Theorem

For any constant c, $E[c] = c$

Theorem

For any random variable W, $E[W + b] = E[W] + b$

Theorem

For any constant c and random variable X,

$$E[cX] = cE[X]$$

Theorem

For any X and Y that are random variables that have finite expectations

$$E[X + Y] = E[X] + E[Y]$$

Expected Value of a Function of a Random Variable

Theorem

Suppose X is a random variable with pdf $p_X(k)$. Let $g(X)$ be any function of X. Then the expected value of the random variable $g(X)$ is given by

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)dx$$

Example

Suppose $X = f_X(x)$ is the uniform density function defined on the range [0,1]. Then $f_X(x) = 1$. Suppose $g(x) = 1/5x$.

$$\begin{aligned} E[g(X)] &= \int_0^1 \frac{1}{5}x \cdot f_X(x)dx \\ &= \frac{1}{5}[\frac{1}{2}x^2]_0^1 = \mathbf{\frac{1}{10}} \end{aligned}$$

Example

The Zetas want to burn their gang-signs into a wall. Suppose the amount of fuel in the torch, Y, is a random variable with pdf

$$f_Y(y) = 3y^2, 0 < y < 1.$$

In the past they have been able to burn a circle whose radius is 20 times the size of y. How much area can they expect to burn?

Answer

By problem setup, $g(Y) = 20\pi Y^2$

$$\begin{aligned} E[g(Y)] &= \int_0^1 20\pi y^2 \cdot 3y^2 dy = 60\pi \int_0^1 y^4 \\ &= \frac{60\pi y^5}{5}\Big|_0^1 \\ &= 12\pi ft^2 \end{aligned}$$

Variance of Discrete Random Variables

Definition

The *variance* of a random variable is the expected value of its squared deviations from μ. If X is discrete with pdf $p_X(k)$,

$$VAR[X] = \sigma^2 = E[(X-\mu)^2] = \sum_{\forall k} (k-\mu)^2 \cdot p_X(k).$$

Note that $\mu = E[X]$. Also note that the standard deviation σ is the square root of the variance.

$$VAR[X] = E[(X - E[X])^2]$$

Just as the expected value has a clear analogy to the "center of balance" of a distribution, the variance is analogous to the "moment of inertia."

Imagine placing two distributions on a turntable and pushing on each with identical force 6" from the center. A distribution with low variance would spin faster, and a distribution with high variance would spin slower.

Example of Discrete RV Variance

Example

An urn contains five chips, two red and three white. Two chips are drawn at random, *with replacement*. Let X denote the number of red chips drawn. What is VAR[X]?

Answer

X is binomial, $n = 2$ and $p = \frac{2}{5}$.

$$
\begin{aligned}
E[X] &= np = (2)(2/5) = 4/5 \\
VAR[X] &= \sum (x - E[X])^2 f(x) = \sum_{i=0}^{2} (i - .8)^2 \binom{2}{i} \left(\frac{2}{5}\right)^i \left(\frac{3}{5}\right)^{2-i} \\
&= (0 - .8)^2 \cdot \binom{2}{0} (2/5)^0 (3/5)^2 + (1 - .8)^2 \cdot \binom{2}{1} (2/5)^1 (3/5)^1 + (2 - .8)^2 \binom{2}{2} (2/5)^2 (3/5)^0 \\
&= (16/25)(1)(9/25) + (1/25)(2)(6/25) + (36/25)(1)(4/25) = \mathbf{\frac{12}{25}}
\end{aligned}
$$

Variance of Continuous Random Variables

Definition

Let X be a real-valued random variable with density function $f_X(x)$. Then, the *variance*, σ^2, is defined by

$$\begin{aligned}\sigma^2 &= E[(X-\mu)^2] \\ &= \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx\end{aligned}$$

A Variance Identity

Theorem

Let W be any random variable (discrete or continuous), having mean μ and for which $E[W^2]$ is finite. Then,

$$VAR[W] = E[(W - E[W])^2] = E[W^2] - E[W]^2$$

Proof.

$$\begin{aligned}
VAR[W] &= E[(W - E[W])^2] \\
&= E[W^2 - 2WE[W] + E[W]^2] \\
&= E[W^2] - E[2XE[W]] + E[E[W]^2] \\
&= E[W^2] - E[2W]E[W] + E[W]^2 \\
&= E[W^2] - 2E[W]E[W] + E[W]^2 \\
&= E[W^2] - E[W]^2
\end{aligned}$$

An Example

Example

Given the following pdf, what is the variance?

$$f(x) = 3(1-x)^2, 0 < x < 1$$

Answer

$$\begin{aligned} E[X] &= \int_0^1 x \cdot 3(1-x)^2 dx = 3\int_0^1 x - 2y^2 + y^3 = 3 \\ E[X^2] &= \int_0^1 x^2 \cdot 3(1-x)^2 dx = 3\int_0^1 y^2 - 2y^3 + y^4 dx = \frac{1}{10} \\ VAR[X] &= E[X^2] - E[X]^2 = \frac{3}{80} \end{aligned}$$

Variance Identities

Theorem

For and random variable X and constant c

$$VAR[cX] = c^2 VAR[X]$$

$$VAR[X + c] = VAR[X]$$

Proof.

For ease, let $E[X] = \mu$. Then $E[cX] = c\mu$, and

$$\begin{aligned} V[cX] &= E[(cX - c\mu)^2] = E[c^2(X - \mu)^2] \\ &= c^2 E[(X - \mu)^2] \\ &= c^2 VAR[X] \end{aligned}$$

□

Variance Identities

Theorem

Let X and Y be two independent random variables. Then,

$$VAR[X+Y] = VAR[X] + VAR[Y]$$

Proof.

Let E[X] = a and E[Y] = b.

$$\begin{aligned} VAR[X+Y] &= E[(X+Y)^2] - (a+b)^2 \\ &= E[X^2] + 2E[XY] + E[Y^2] - a^2 - 2ab - b^2 \\ &= E[X^2] - a^2 + E[Y^2] - b^2 = VAR[X] + VAR[Y] \end{aligned}$$

□

Discrete Joint pdfs

Definition

Suppose S is a discrete sample space on which two random variables, X and Y, are define. The *joint probability density function* of X and Y (called the "joint pdf") is denoted $p_{X,Y}$(x,y), where,

$$p_{X,Y}(x,y) = P(\{s|X(s) = x \text{ and } Y(s) = y\})$$

Example

A polling site has two voting booths. The joint pdf of people in line at the booth at 7am is given by the following table:

Joint pdf Example

Example

Y\X	0	1	2	3
0	0.1	0.2	0	0
1	0.2	0.25	0.05	0
2	0	0.05	0.05	0.25
3	0	0	0.025	0.05

Find the probability that X and Y have exactly one more or less people in line, $P(|X - Y| = 1)$

Answer

$$\begin{aligned} P(|X - Y| = 1) &= \sum \sum_{|x-y|=1} p_{X,Y}(x, y) \\ &= 0.2 + 0.2 + 0.05 + 0.05 + 0.025 + 0.025 \\ &= 0.55 \end{aligned}$$

Joint pdf Theorem

Theorem

Suppose that $p_{X,Y}(x,y)$ is the joint pdf of the discrete random variables X and Y. Then,

$$p_X(x) = \sum_{\forall Y} p_{X,Y}(x,y) \text{ and } p_Y(y) = \sum_{\forall X} p_{X,Y}(x,y)$$

Proof.

Note that the collection of sets (Y=y) for all y forms a partition of Ω. Then, the set $(X = x) = (X = x) \cap S = (X = x) \cap \bigcup_{\forall y}(Y = y) = \bigcup_{\forall y}[(X = x \cap (Y = y)]$, and so

$$\begin{aligned} p_X(x) &= P(X = x) = P\left(\bigcup_{\forall y}[(X = x) \cap (Y = y)]\right) \\ &= \sum_{\forall y} P(X = x, Y = y) = \sum_{\forall y} p_{X,Y}(x,y) \end{aligned}$$