EE126: Probability and Random Processes

Lecture 6: Expectation and Variance of Discrete Random Variables

Abhay Parekh

UC Berkeley

February 3, 2011

1. Logistics
2. Review
3. Expected Value and Variance

HW Logistics

1. HW Box marked EE126 in the Student Lounge, 240 Cory.
2. Must be in by end of day Tuesday.
3. All HW must be submitted this way.
4. No late HW will be accepted.

Random Variables

- A random variable is a function from the outcomes of an experiment to the set of real numbers.
- **Discrete rvs:** can take on a finite or countably infinite number of values.

  Continuous rvs: take on an uncountably infinite number of values.

  Example: A point is chosen at random on the line $[0, 1]$.
  - Continuous r.v. a: the value of the point chosen
  - Discrete r.v. b: $b = 1$ if $a \geq 0.5$ and $b = 0$ otherwise.

Probability Mass Functions

Defines the probability for each possible value of a discrete random variable.

Generally, for each possible value x of X:

1. Collect all the possible outcomes that give rise to $\{X = x\}$.
2. Add their probabilities to obtain $p_X(x)$.

For any random variable, X,

$$\sum_x p_x(x) = 1.$$

Common Random Variables

1. Geometric Random Variable **counts the time to first success**.
   The probability of success is p and we keep performing independent trials of an experiment until the first success occurs.

$$p_X(k) = (1-p)^{k-1} p$$

2. Bernoulli RV: **Simple boolean.** $B = 1$ if a coin flips head and $B = 0$ otherwise.
3. Binomial Random Variable: **Counts the number of successes.** $X=$ number of heads in N coin tosses.

$$p_X(k) = \binom{N}{k} p^k (1-p)^{N-k}$$

4. Poisson Random Variable: **Approximates a Binomial** When N is large and p is small an excellent approximation is:

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

   where $\lambda = Np > 0$.

Functions of Random Variables

Given a random variable X and a function $Y = g(X)$:

$$p_Y(y) = \sum_{\{x : g(x) = y\}} p_X(x).$$

Expected Value (Mean)

- When a coin is tossed 3 times, what is the "average" number of heads obtained?
- X: the number of heads obtained. Suppose the experiment is performed N times, and N is large.
- Interpret $p_X(x)$ as the fraction of the experiments for which $X = x$.
- There are about $Np_X(x)$ experiments for which the outcome is x.
- Then the average value is

$$\sum_{i=1}^{3} \frac{iNp_X(i)}{N} = \sum_{0=1}^{3} ip_X(i).$$

- Thus

$$E[X] =$$

Expected Value of a Discrete RV

The expected value of X is:

$$E[X] = \sum_x x p_X(x).$$

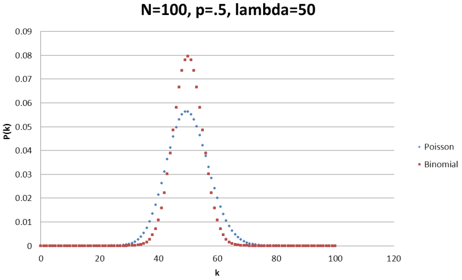
$E[X]$ is like the "center of mass" of the probability mass function.
Also, to be well defined, $\sum_x |x| p_X(x) < \infty$.

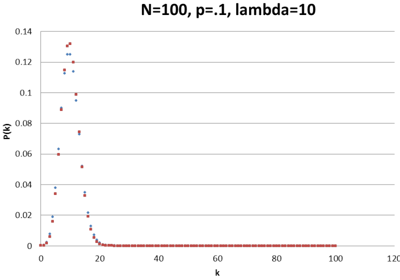
Example: Binomial Random Variable

What is the expected number of H when a coin with P(H) is tossed N times.

X: number of heads. This a binomial random variable.

$E[X] = \sum_{k=0}^{N} k \binom{N}{k} p^k (1-p)^{N-k}$. Plot and guess...





Looks like the answer is Np. But that's just a guess...

Example: Binomial Random Variable

$$E[X] = \sum_{k=0}^{N} k \binom{N}{k} p^k (1-p)^{N-k}.$$

Notice that the first term is zero.

$$\begin{aligned}
E[X] &= Np \sum_{k=1}^{N} k \frac{N!}{k!(N-k)!N} p^{k-1} (1-p)^{N-k} \\
&= Np \sum_{k=1}^{N} \frac{(N-1)!}{(k-1)!(N-k)!} p^{k-1} (1-p)^{N-k} \\
&= Np \sum_{k=1}^{N} \binom{N-1}{k-1} p^{k-1} (1-p)^{N-k} \\
&\quad \text{Let } i := k-1 \\
&= Np \sum_{i=0}^{N-1} \binom{N-1}{i} p^{i} (1-p)^{N-1-i} \\
&= Np
\end{aligned}$$

Fair Game

Game 1: You toss a fair coin repeatedly until you get a heads. If this takes n tosses, you get n dollars. What is the most you should pay to play this game?
W_1: the amount you win.

$$E[W_1] = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2$$

$P(W_1 > 2) = \frac{1}{4}$.

St. Petersburg Paradox

Game 2: You toss a fair coin repeatedly until you get a heads. If this took n tosses, you get 2^n dollars. What is the most you should pay to play this game?
W_2:the expected amount you would win.

$$E[W_2] = \sum_{i}^{\infty} \frac{2^i}{2^i}$$

So you should definitely pay all the money you have (and can borrow). Or should you? Suppose you pay 2^n.
$P(W_2 > 2^n) = \frac{1}{2^n}$.

$E[g(x)]$

Suppose we are interested in $Y = X^2 + 3$, e.g. if the experiment yields 2 heads, the value of Y is 7. Then what is $E[Y]$?

$$E[Y] = \sum_y y p_Y(y)$$

1. The possible values of Y are $3, 4, 7, 12$.
2. Thus
$$p_Y(y) = \begin{cases} p_X(0) = \frac{1}{8}, & \text{y=3;} \\ p_X(1) = \frac{3}{8}, & \text{y=4;} \\ p_X(2) = \frac{3}{8}, & \text{y=7;} \\ p_X(3) = \frac{1}{8}, & \text{y=12.} \end{cases}$$
3. $$E[Y] = 3 * \frac{1}{8} + 4 * \frac{3}{8} + 7 * \frac{3}{8} + 12 * \frac{1}{8} = 6$$

Can $E[Y]$ can be evaluated in terms of $p_X(x)$? Do we really need to find $p_Y(y)$?

Expected Value Rule for Functions of RVs

Let X be a rv with PMF p_X and let $g(X)$ be a function of X. Then

$$E[g(X)] = \sum_x g(x)p_X(x).$$

Start with the basic outcomes:

Ω:	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
X:	3	2	2	1	2	1	1	0
$Y = g(X)$:	g(3)	g(2)	g(2)	g(1)	g(2)	g(1)	g(1)	g(0)

$$E[Y] = g(0)p_X(0) + g(1)p_X(1) + g(2)p_X(2) + g(3)p_X(3)$$

$$E[Y] = \sum_x g(x)p_X(x).$$

Examples

Let X be a rv with PMF p_X and let $g(X)$ be a function of X. Then

$$E[g(X)] = \sum_x g(x)p_X(x).$$

- $E[a] =$
- $E[aX] =$

Variance and Standard Deviation

$p_A(a) = \frac{1}{1001}$ if $a = -500, -499, ..., 0, ...500$
$p_B(b) = 1$ if $b = 0$.
$E[A] = E[B] = 0$ but A is much more *dispersed* around its mean.
How should we capture this dispersion?
What about measuring it by $X - E[X]$? $(X - E[X])^2$ is better.

- Variance: $\mathrm{var}(X) = \sigma_X^2 = E[(X - E[X])^2]$ units are mean^2.
- Standard Deviation $= \sigma_X = \sqrt{\mathrm{var}(X)}$

Letting $g(X) = (X - E[X])^2$:

$$\mathrm{var}(X) = \sum_x (x - E[X])^2 p_X(x).$$

1. $\mathrm{var}(X) \geq 0$.
2. $var(a) = 0$.

Calculating Variance I

Example: X: number of heads in three tosses of a fair coin.
$E[X] = 1.5$;

$$\mathrm{var}(X) = \sum_x (x - E[X])^2 p_X(x).$$

$$\begin{aligned}\mathrm{var}(X) &= (0 - \frac{3}{2})^2 \frac{1}{8} + (1 - \frac{3}{2})^2 \frac{3}{8} + (2 - \frac{3}{2})^2 \frac{3}{2} + (3 - \frac{3}{2})^2 \frac{1}{8} \\ &= \frac{3}{4} \\ \sigma_X &= \sqrt{3}/2\end{aligned}$$

Calculating Variance II

Convenient formula:

$$\mathrm{var}(X) = E[X^2] - E[X]^2$$

Example: X: number of heads in three tosses of a fair coin.

$$E[X^2] = \frac{3}{8} + \frac{4\;3}{8} + \frac{9}{8} = 3$$

$$\mathrm{var}(X) = 3 - (1.5)^2 = 0.75$$

Fair Game

Game 1: You toss a fair coin repeatedly until you get a heads. If this takes n tosses, you get n dollars. What is the most you should pay to play this game?
W_1: the amount you win.

$$E[W_1] = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2$$

$P(W_1 > 2) = \frac{1}{4}$.

$$\mathrm{var}(W_1) = E[W_1^2] - E[X]^2$$

$$\mathrm{var}(W_1) = 2.$$

Probability that that the winnings are $\pm$ one standard deviation from the mean:

$$P(2 - \sqrt{2} \leq W_1 \leq 2 + \sqrt{2}) = P(W_1 \leq 3) = \frac{7}{8}$$

Properties of Expectation and Variance

$$Y = aX + b$$

1. $$\begin{aligned} E[Y] &= \sum_x (ax+b)p_X(x) = a\sum_x x p_X(x) + b\sum_x p_X(x) \\ &= aE[X] + b \end{aligned}$$
2. $\mathrm{var}(Y) =$

Example: Celsius and Fahrenheit

A city has a mean and standard deviation equal to 10 degrees Celsius. A "typical" day is the average + or - one standard deviation from the mean. What is the temperature of a typical day in Fahrenheit?

C: city temperature in Celsius. $E[C] = 10$, $\mathrm{var}(C) = 100$.
$F = 32 + 1.8C$.

$$E[F] = E[32 + 1.8C] = 32 + 1.8E[C] = 32 + 18 = 50$$

$$\mathrm{var}(F) = \mathrm{var}(32 + 1.8C) = 2.24\mathrm{var}(C) = 224$$

$$\sigma_F = 18$$

A typical day is 50 ± 18 Fahrenheit.

Example: Normalizing a Random Variable

We are given values from a random variable X with $E[X] = \mu$ and variance σ^2. We want to convert this to a random variable X^* such that $E[X^*] = 0$ and $\sigma^2_{X^*} = 1$.

1. $X_1^* = X - \mu$ has mean 0
2. $X_2^* = \frac{X}{\sigma}$ has variance $\frac{1}{\sigma^2}\mathrm{var}(X) = 1$.
3. $X^* = \frac{X-\mu}{\sigma}$ has mean 0 and variance 1.

Example: Average Speed v/s Average Time

If the weather is bad in Berkeley (which happens with probability 0.4), Johnny will leave early for work (25 miles away) and miss rush hour. His average speed is then 50mph. Otherwise he leaves late, hits the rush hour and his average speed is 25mph. What is the mean time he takes to get to work?

V: his speed to work; T: his time to get to work; $T = \frac{25}{V}$.

$$E[V] = (.4)50 + (.6)25 = 35$$

Quite natural to say that $E[T] = \frac{25}{35} = \frac{5}{7}$. But this is WRONG.

Why? $E[\frac{25}{V}] \neq \frac{25}{E[V]}$.

A correct way:

$$\begin{aligned} E[T] &= \sum_v \frac{25}{v} p_V(v) \\ &= \frac{25}{50}.4 + \frac{25}{25}.6 \\ &= .2 + .6 = .8 \end{aligned}$$

$E[X]$ for non negative random variables

Suppose X only takes on non negative values. Then

$$E[X] = \sum_{i=1}^{\infty} P(X \geq k).$$

$$\begin{aligned}
\sum_{i=1}^{\infty} P(X \geq k) &= \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} p_X(j) \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{i} p_X(i) \\
&= \sum_{i=1}^{\infty} i p_X(i)
\end{aligned}$$

Example: X is a geometric random variable.

$$E[X] = \sum_{i=1}^{\infty} i(1-p)^{i-1} p = \sum_{i=1}^{\infty} (1-p)^{i-1}$$

$$E[X] = \frac{1}{p}$$