EE126: Probability and Random Processes

Lecture 8: Conditioning and Independence of Discrete Random Variables

Abhay Parekh

UC Berkeley

February 10, 2011

Dealing with Multiple Random Variables

Multiple random variables are like multiple **events**.

$$P_{X,Y}(x,y) = P(X = x, Y = y) = P(\{X = x\} \cap \{Y = y\}).$$

Given rvs X and Y define $Z = g(X, Y)$ E.g. $Z = X + Y$.
Deal with this just like we did with functions of a single rv:

$$p_Z(z) = \sum_{x:g(x,y)=z} p_{XY}(x,y).$$

By extension of the result for a single rv:

$$E[g(X,Y)] = \sum_{x,y} g(x,y)p_{XY}(x,y).$$

Expected values of Sums of Random Variables

Let $g(X, Y) = X + Y$

$$E[aX + bY + c] = aE[X] + bE[Y] + c.$$

Using indicator random variables can often be helpful: Express X as the sum of N boolean random variables $X_1, ...X_N$ and then find $\sum_{i=1}^{N} E[X_i]$.

Conditioning of Random Variables

Very similar to conditioning of events.

1. Conditioning on an Event:

$$P(X = k|A) = \frac{P(X = k \cap A)}{P(A)}.$$

2. Conditioning on another Random Variable:

$$P(X = k|Y = m) = \frac{P(X = k, Y = m)}{P(Y = m)} = p_{X|Y}(k|m).$$

$p_{X|Y}(x|y)$ is called the conditional pmf of X given Y.
Rewriting:

$$p_{X|Y}(x|y) = \frac{p_{XY}(x, y)}{p_Y(y)}$$

Conditional distributions

1. $X|Y$ is a random variable: $\sum_x p_{X|Y}(x|y) = \sum_x \frac{p_{X,Y}(x,y)}{p_Y(y)} = 1$
2. Multiplication Rule:
   $p_{X,Y}(x,y) = p_X(x)p_{Y|X}(y|x) = p_Y(y)p_{X|Y}(x|y)$.
3. Total Probability Theorem: If $A_1, \ldots, A_N$ partition the sample space and $P(A_i) > 0$ for each i:

$$p_X(x) = \sum_{i=1}^{n} P(A_i)p_{X|A_i}(x).$$

4. Nothing special about just two random variables, naturally extends to more.

Example: Three tosses of a fair coin

X: number of heads; Y: indicator function on whether second toss is a head. ($Y = 1$ if second toss is a heads and $Y = 0$ otherwise.)

$p_{X|Y=0}(x) = \frac{p_{X,Y}(x,0)}{p_Y(0)} = 2p_{X,Y}(x,0)$

	0	1
0	$\frac{1}{8}$	0
1	$\frac{2}{8}$	$\frac{1}{8}$
2	$\frac{1}{8}$	$\frac{2}{8}$
3	0	$\frac{1}{8}$

$$p_{X|Y=0}(x) = \begin{cases} 1/4, & \text{x=0;} \\ 1/2, & \text{x=1;} \\ 1/4, & \text{x=2;} \\ 0, & \text{x=3.} \end{cases}$$

Example: Eleven Urns

There are 11 identical looking urns $u = 0, 1, 2, ..., 10$ each containing 10 balls. Urn u has u black balls and $10 - u$ white balls. You pick an urn at random and draw 10 balls with replacement and obtain 3 black balls. What is the probability that you selected urn u?
You are estimating a probability distribution based on 10 observations.
u: urn selected; Find $P(u|n_B, N)$, $n_B = 3$, $N = 10$.

$$P(u|n_B, N) = \frac{P(u, n_B|N)}{P(n_B|N)} = \frac{P(n_B|u, N)P(u)}{P(n_B|N)}$$

$P(u) = \frac{1}{11}$, $\quad P(n_B|u, N) = \binom{N}{n_B}(u/10)^{n_B}(1 - u/10)^{N-n_B}$

$$P(n_B|N) = \sum_u P(u, N_B|N) = \sum_u P(u)P(n_B|u, N)$$

Now, specializing to the known values, we get
$P(n_B = 3|N = 10) = 0.083$.

Example: Eleven Urns

$$P(u|n_B, N) = \frac{P(u, n_B|N)}{P(n_B|N)} = \frac{P(n_B|u, N)P(u)}{P(n_B|N)}$$

$$P(u|n_B = 3, N = 10) = \frac{1}{0.083}\frac{1}{11}\binom{10}{3}\frac{u}{10}^3(1 - \frac{u}{10})^7$$

u	$P(u\|n_B = 3, N = 10)$
0	0
1	0.063
2	.22
3	0.29
4	0.24
5	0.13
6	0.047
7	0.0099
8	0.00086
9	0.0000096
10	0

Relationship to coin tosses...

Seats on a Plane

Passengers on a full flight of N seats sit at random. What is the probability that no passenger is their assigned seat?
Z_n: event that no one is their assigned seat in an n seat plane.
$X_1 = 1$ if passenger 1 ends up in seat 1 and $X_1 = 0$ otherwise.

$$P(Z_n) = P(Z_n|X_1 = 1)P(X_1 = 1) + P(Z_n|X_1 = 0)P(X_1 = 0) = P(Z_n|X_1 = 0)\frac{n-}{n}$$

$P(Z_n|X_1 = 0)$: prob that none of n-1 passengers takes their assigned seats, when one passenger (the "extra") couldn't possibly take their assigned seat.
Two mutually exclusive possibilities: Either extra person chooses seat 1 (A) or he does not (A^c).

$$P(Z_n|X_1 = 0) = Z_{n-2}P(A) + P(Z_n, A^c|X_1) = Z_{n-2}\frac{1}{n-1} + Z_{n-1}$$

Seats on a Plane

Two mutually exclusive possibilities: Either extra person chooses seat 1 (A) or he does not (A^c).

$$P(Z_n|X_1 = 0) = P(Z_{n-2})P(A)+P(Z_n, A^c|X_1) = P(Z_{n-2})\frac{1}{n-1}+P(Z_{n-1}$$

$$P(Z_n) = \frac{n-1}{n}P(Z_{n-1}) + \frac{1}{n}P(Z_{n-2})$$

$$P(Z_n) - P(Z_{n-1}) = -\frac{1}{n}(P(Z_{n-1}) - P(Z_{n-2}))$$

$Z_1 = 0, Z_2 = \frac{1}{2}$ so $Z_3 =$, $Z_4 =$

$$P(Z_n) = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} + ... + \frac{(-1)^n}{n!}$$

Conditional Expectation

Recognize that $X|Y = y$ and $X|A$ (where A is an event with $P(A) > 0$) are random variables:

1. If A is an event such that $P(A) > 0$: $E[X|A] = \sum_x x p_{X|A}(x)$
2. $E[X|Y = y] = \sum_x x p_{X|Y=y}(x|y)$
3. $E[g(X)|A] = \sum_x g(x) p_{X|A}(x)$
4. Total Expectation Theorem: If $A_1, ..., A_N$ partition the sample space and $P(A_i) > 0$ for each i:
   - $$E[X] = \sum_{i=1}^{n} P(A_i) E[X|A_i].$$
   - Given B such $P(A_i \cap B) > 0$ for each i:
     $$E[X|B] = \sum_{i=1}^{n} P(A_i) E[X|A_i \cap B].$$
5. Nothing special about just two random variables, naturally extends to more.

Example: Three tosses of a fair coin

X: number of heads; Y: indicator function on whether second toss is a head. ($Y = 1$ if second toss is a heads and $Y = 0$ otherwise.)

$$E[X + Y|X \geq 1] = E[X|X \geq 1] + E[Y|X \geq 1]$$

	0	1
0	$\frac{1}{8}$	0
1	$\frac{2}{8}$	$\frac{1}{8}$
2	$\frac{1}{8}$	$\frac{2}{8}$
3	0	$\frac{1}{8}$

$$p_{X|X\geq 1}(x) = \frac{p_X(x)}{P(X \geq 1)}, \quad x = 1, 2, 3;$$

$$p_{Y|X\geq 1}(x) = \frac{P(\{Y = y\} \cap \{X \geq 1\})}{P(X \geq 1)}, \quad y = 0, 1;$$

Mean of the Geometric Random Variable

X counts the number of coin tosses to the first head (prob of heads is p). Suppose we have tossed the coin k times with no success. Does that affect the future?

The Geometric RV is **memoryless**:

I.e:

$$p_{X|X\geq k}(k+m) = \frac{p_X(k+m)}{\sum_{i=k}^{\infty} p(1-p)^i} = p(1-p)^{m-1} = p_X(m).$$

Thus $E[X|X>1] = 1 + E[X]$. Let's use this:

$$\begin{aligned} E[X] &= P(X=1)E[X|X=1] + P(X>1)E[X|X>1] \\ &= p + (1-p)(1+E[X]) \\ &= \frac{1}{p} \end{aligned}$$

Variance of the Geometric Random Variable

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

$$\begin{aligned}
E[X^2] &= P(X=1)E[X^2|X=1] + P(X>1)E[X^2|X>1] \\
&= p + (1-p)E[(1+X)^2] \\
&= p + (1-p)E[X^2] + 1 - p + 2(1-p)\frac{1}{p} \\
&= (1-p)E[X^2] + \frac{2(1-p)}{p} \\
&= \frac{2(1-p)}{p^2}
\end{aligned}$$

$$var(X) = E[X^2] - E[X]^2 = \frac{2(1-p)}{p^2} - \frac{1}{p^2}$$

$$var(X) = \frac{1-p}{p^2}$$

Example: Problem of Joint Lives

2m people form couples today. 50 years from now, the probability of any person being alive is p. Now suppose there are A people alive 50 years from now and let S be the number of couples for which both partners are still alive. Find $E[S|A=a]$.
If we can find $p_S(S|A=a)$ we can find $\sum_s sp_S(S|A=a)$.

$$p_S(s|A=a) = \frac{P(S=s \cap A=a)}{P(A=a)}$$

$$P(A=a) = \binom{2m}{a} p^a (1-p)^{2m-a}$$

But finding the numerator is hard!
New approach exploits linearity of expectation.

Example: Problem of Joint Lives

Think of the people lined up. $X_i = 1$ if the first person in couple i is alive in 50 years and $X_i = 0$ otherwise. $Y_i = 1$ if the second person in couple i is alive in 50 years and $Y_i = 0$ otherwise. Now $S = \sum_{i=1}^{m} X_i Y_i$.

$$E[S|A=a] = E[\sum_{i=1}^{m} X_i Y_i | A=a] = \sum_{i=1}^{m} E[X_i Y_i | A=a].$$

Each of the terms has the same value...

$$\begin{aligned} E[S|A=a] &= mE[X_1 Y_1 | A=a] \\ &= mE[Y_1 = 1 | X_1 = 1, A=a] P(X_1 = 1 | A=a) \\ &\quad \text{but } Y_1 \text{ is a boolean rv} \\ &= mP(Y_1 = 1 | X_1 = 1, A=a) P(X_1 = 1 | A=a) \\ &= m \frac{a-1}{2m-1} \frac{a}{2m} = \frac{a(a-1)}{2m(2m-1)} \end{aligned}$$

Why independent of p?