EE126: Probability and Random Processes

Lecture 7: Joint PMFs and Conditioning of Discrete Random Variables

Abhay Parekh

UC Berkeley

February 8, 2011

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.

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, $\sum_x |x| p_X(x) < \infty$.
Suppose X only takes on non negative values. Then

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

Computing $E[g(x)]$

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

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

- $E[a] = a$
- $E[aX] = aE[X]$
- $E[g(X)] \geq g(E[X])$.

Variance and Standard Deviation

- 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. $var(a) = 0$.

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

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. $\mathrm{var}(Y) = a^2 \mathrm{var}(X)$

Dealing with Multiple Random Variables

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

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

Example: X : #heads in three tosses of a coin, Y: $Y = 0$ if no Head in the second toss and $Y = 1$ ow.

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

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

$p_X(x)$ and $p_Y(y)$ are called marginal PMFs and $p_{XY}(x,y)$ is the joint PMF.

Functions of multiple random variables

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:

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

Previous Example: $Z = X + Y$:

$$Pr(X + Y = k) = \sum_i Pr(X = i, Y = k - i).$$

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

Z	0	1	2	3	4
$P_Z(z)$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

$E[g(X,Y)]$

By extension of the result for a single rv:

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

Why is

$$\sum_{x,y} g(x,y)p_{XY}(x,y) = \sum_{k} kP(g(X,Y)=k)?$$

$$\begin{aligned}
\sum_k kP(g(X,Y)=k) &= \sum_k k \sum_{x,y:g(x,y)=k} p_{XY}(x,y) \\
&= \sum_k \sum_{x,y:g(x,y)=k} k p_{XY}(x,y) \\
&= \sum_k \sum_{x,y:g(x,y)=k} g(x,y) p_{XY}(x,y) \\
&\quad \text{no double counting of } p_{XY}(x,y) \text{ since } g(x,y) \text{ is a function} \\
&= \sum_{x,y} g(x,y) p_{XY}(x,y)
\end{aligned}$$

Expected values of Sums of Random Variables

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

$$\begin{aligned} E[X+Y] &= \sum_{x,y} g(x,y) p_{XY}(x,y) \\ &= \sum_{x,y} x p_{XY}(x,y) + \sum_{x,y} y p_{XY}(x,y) \\ &= E[x] + E[Y] \end{aligned}$$

Remarkable result. True even if X and Y are dependent!
Also,

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

Previous Example: $E[X] =\quad$, $E[Y] =\qquad E[X+Y] =\quad$. Matches.

Example: Binomial Random Variable

Given N trials with p as the probability of success, we know that the number of successes $X = \sum_{i=1}^{N} B_i$ where B_i is a Bernouli rv with prob of success p. Then

$$E[X] = E[\sum_i B_i] = \sum_i E[B_i] = Np.$$

Seats on a Plane

Even though the passengers on a full flight have assigned seats, they disregard this and sit at random. Let X be the number of passengers who still end up in their assigned seats. What is $E[X]$?
Suppose there are N passengers, N seats. Finding $p_X(x)$ is not easy....
Way around this. Let $X_i = 1$ if passenger i ends up in his assigned seat and $X_i = 0$ otherwise.

$$X = X_1 + X_2 + X_3 + \ldots + X_N$$

$$E[X] = \sum_{i=1}^{N} E[X_i]$$

But X_i is a Bernouli rv with probability of success $= \frac{1}{N}$.

$$E[X_i] = (1)\frac{1}{N} + 0\frac{N-1}{N} = \frac{1}{N}$$

Thus

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

Independent of N. Somewhat surprising...

Example: Sampling the whole range

N balls, each a different color. We sample with replacement. What is the expected number of trials until we've drawn at least one ball of each color?

Define a "successful trial" to be one in which we have drawn a new color. Once we have N successful trials we are done. V_r: number of trials to the r^{th} success.

$V_1 = 1$. Define $X_1 = 1$ and $X_k = V_k - V_{k-1}$, $k = 2, 3, ..., N$ is a geometric random variable with $p_k = \frac{N-k+1}{N}$.

$$V_N = X_1 + \ldots + X_N$$

$$\begin{aligned} E[V_N] &= E[X_1] + \ldots + E[X_N] \\ &= N\left(\frac{1}{N} + \frac{1}{N-1} + \ldots + \frac{1}{1}\right) \end{aligned}$$

For $E[V_6] = 14.7$

Example: Knockout Task Sequencing

There are N tasks. Completing task i correctly gives you a reward of v_i dollars and the right to continue to the next task. Completing it incorrectly results in immediate disqualification. You can keep your winnings. If p_i is the the prob you will complete task i correctly. What sequence of the task maximizes your expected reward?

1. Pick a sequence L. Say it is numbered so that $L = (1, 2, 3, ..., n)$ (WLOG).
2. Let $R(L)$ be the reward when sequence L is followed.
3. For some k create a new sequence L_k that interchanges the order of task k and $k+1$ in sequence L.
4. $$E[R(L)] = \sum_k k p_{R(L)}(k) = \sum_{k=1}^{\infty} P(R(L) \geq k)$$

Example: Knockout Task Sequencing

$$E[R(L)] = \sum_{k=1}^{\infty} P(R(L) \geq k)$$

$$P(R(L) \geq k) = \begin{cases} p_1, & k = 1, 2, ..., v_1; \\ p_1 p_2, & k = v_1 + 1, ..., v_1 + v_2; \\ p_1 p_2 p_3, & k = v_1 + v_2 + 1, ..., v_1 + v_2 + v_3; \\ p_1 p_2 ... p_n, & k = \sum_{i=1}^{n-1} v_i + 1, ..., \sum_{i=1}^{n} v_i. \end{cases}$$

$$E[R(L)] = p_1 v_1 + p_1 p_2 v_2 + p_1 p_2 p_3 v_3 + \ldots p_1 p_2 ... p_n v_n.$$

Example: Knockout Task Sequencing

4.

$$E[R(L)] = \underbrace{p_1 v_1 + p_1 p_2 v_2 + \ldots + p_{k-1} v_{k-1}}_{A} + \underbrace{p_1 \ldots p_k v_k + p_1 \ldots p_{k+1} v_{k+1}}_{B} + \underbrace{p_1 \ldots p_{k+2} v_{k+2} + \ldots p_1 \ldots p_N v_N}_{C}$$

5.

$$E[R(L_k)] = A + \underbrace{p_1 p_2 \ldots p_{k-1} p_{k+1} v_{k+1} + p_1 p_2 \ldots p_{k-1} p_{k+1} p_k v_k}_{B_k} + C$$

6.

$$E[R(L_k)] - E[R(L)] = B_k - B = p_1 .. p_{k-1}(1 - p_{k-1})(1 - p_{k+1})(w_{k+1} - w_k)$$

where $w_i = \frac{p_i v_i}{(1 - p_i)}$.

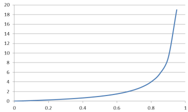
Example: Knockout Task Sequencing

7.
$$E[R(L_k)] \geq E[R(L)] \iff w_{k+1} - w_k \geq 0$$

8. There must be at least one optimal sequence with non decreasing weights.

9. If all the w_i are unique there is only one sequence. If not and $w_i = w_{i+1}$ and interchanging the pair does not change the reward. Thus any two non decreasing sequences can reached reached from each other by repeated interchanges and they all have the same expected reward.

10. To maximize the expected reward order the tasks by non-decreasing:
$$w_i = \frac{p_i v_i}{(1 - p_i)}.$$



Variance?

$$var(g(X, Y)) = E[(g(X, Y) - E[g(X, Y)])^2] = E[(g(X, Y)^2] - E[g(X, Y)$$

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

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:

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

$p_{X|Y}(x|y)$ 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 $A_1, ..., A_N$ 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.