Corollary

If $X_1, \ldots, X_k$ are independent, then

$$(1.2.3) \qquad E(X_1 \cdots X_k) = E(X_1) \cdots E(X_k)$$

and

$$(1.2.4) \qquad Var(a_1 X_1 + \cdots + a_k X_k) = a_1^2 Var(X_1) + \cdots + a_k^2 Var(X_k)$$

for numbers $a_1, \ldots, a_k$, provided $Var(X_i) < \infty$, $1 \le i \le k$.

#2, 1/24

§1.3 Conditional Probability

We will be greatly concerned with evaluating statements like "What is the probability of event A given event B has occurred?"

Suppose we conduct an experiment N times and we observe the occurrences of two events A, B. We are only interested in outcomes for which B occurs and disregard others. In this smaller set, the proportion of times A occurs is

$$\frac{N_{A \cap B}}{N_B}$$

since B occurs in both. We write

$$\frac{N_{A \cap B}}{N_B} = \frac{N_{A \cap B}/N}{N_B/N}.$$

This suggests that the probability that A occurs given B has occurred is

$$\frac{P(A \cap B)}{P(B)}.$$

Definition 1.3.1

Let B be an event with $P(B) > 0$. The conditional probability that event A occurs given that B occurred is

$$(1.3.1) \quad P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

This is undefined for events B with $P(B) = 0$.

In stochastic processes, we usually use (1.3.1) in the form

$$(1.3.2) \quad P(A \cap B) = P(A|B)\, P(B)$$

We say "the probability of A given B" and "the probability of A conditioned on B".

Example 1.3.1

Roll a dice and observe a number. Let $A = \{\text{odd outcome}\}$ and $B = \{\text{at least } 4\}$.

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(\{5\})}{P(\{4,5,6\})} = \frac{1/6}{1/2} = 1/3$$

Conditional probability has many important properties. The first standard result says that it is indeed a probability.

<u>Theorem 1.3.1</u>

Let B be a fixed event in a sample space S. Then,

(1) $0 \le P(A|B) \le 1$, for any event A

(2) $P(S|B) = 1$

(3) If $\{A_1, A_2, \ldots\}$ is a sequence of pairwise disjoint events

$$P\left(\bigcup_{k=1}^{\infty} A_k \mid B\right) = \sum_{k=1}^{\infty} P(A_k|B)$$

Since $P(B|B) = 1$, we can think of conditioning on B as choosing a new sample space.

Conditional probability gives a test for independence.

<u>Theorem 1.3.2</u>

If $P(A|B)$ is defined, then A and B are independent if and only if $P(A) = P(A|B)$.

Example 1.3.2

Choose a card at random from a full deck. Let

$$A = \{\text{the card is an ace}\}$$

$$B = \{\text{the card is a heart}\}$$

$P(A) = \frac{4}{52}$, $P(B) = \frac{13}{52} = \frac{1}{4}$, $P(A \cap B) = \frac{1}{52}$, and

$$P(A \cap B) = P(A)P(B)$$

so A and B are independent. We see that $P(A|B) = \frac{1}{13}$

Finally, the basis for much of the analysis in this course is

Theorem 1.3.3 Law of Total Probability

Let $\{B_1, B_2, \dots\}$ be a sequence of events such that

(1) $P(B_i) > 0$, all i

(2) $B_i \cap B_j = \phi$, $i \neq j$

(3) S = sample space $= \bigcup_{j=1}^{\infty} B_j$

For any event A,

$$(1.3.3) \qquad P(A) = \sum_{k=1}^{\infty} P(A|B_k)P(B_k)$$

Proof

Since the $\{B_j\}$ are disjoint

$$A = A \cap S = \bigcup_{k=1}^{\infty} A \cap B_k$$

where the sets $\{A \cap B_k\}$ are disjoint. So

$$P(A) = \sum_{k=1}^{\infty} P(A \cap B_k) = \sum_{k=1}^{\infty} P(A \mid B_k) \cdot P(B_k)$$

Definition 1.3.2

Let X, Y be nonnegative, integer valued random variables. The joint probability mass function is

$$P_{X,Y}(j,k) = P_{XY}(j,k) = P(X = j, Y = k) \quad , j,k = 0,1,2,\ldots$$

The marginal probability mass functions are

$$P_X(j) = \sum_{k=0}^{\infty} P_{X,Y}(j,k)$$

and

$$P_Y(k) = \sum_{j=0}^{\infty} P_{X,Y}(j,k)$$

Definition 1.3.3

Let X, Y be nonnegative, integer valued random variables. The <u>conditional</u> <u>probability</u> <u>mass</u> <u>function</u> of X given $Y=k$ is

$$(1.3.4) \qquad P_{X|Y}(j|k) = \frac{P(X=j,\ Y=k)}{P(Y=k)}$$

when $P(Y=k)>0$ and is undefined otherwise.

Theorem 1.3.4

If X, Y are nonnegative, integer valued random variables,

$$(1.3.5) \qquad P_{X|Y}(j|k) = \frac{P_{X,Y}(j,k)}{P_Y(k)} \quad \text{when } P_Y(k)>0.$$

Theorem 1.3.5

$P_{X|Y}(x|k)$ is a probability mass function in x for each fixed k, i.e.

(1) $0 \le P_{X|Y}(j|k) \le 1$, $j,k = 0,1,2,\ldots$

(2) $\sum_{j=0}^{\infty} P_{X|Y}(j|k) = 1$, all k.

Theorem 1.3.6 Law of Total Probability

If X, Y are nonnegative, integer valued random variables,

$$P_X(j) = \sum_{k=0}^{\infty} P_{X|Y}(j|k)\, P_Y(k) \qquad (1.3.6)$$

Proof

Exercise

Example 1.3.3

Let X have a binomial distribution with parameters p, N, where N is a random variable with binomial distribution with parameters θ, m. What is the marginal distribution of X?

We are given the c.p.m.f.

$$P_{X|N}(k|n) = \binom{n}{k} p^k (1-p)^{n-k} \quad , k = 0, 1, 2 \ldots, n$$

(be sure you understand this!)

and the m.p.m.f. for N:

$$P_N(n) = \binom{m}{n} \theta^n (1-\theta)^{m-n} \quad , \quad n = 0, 1, \ldots, m.$$

Then

- $$P(\mathbb{X}=k) = \sum_{n=k}^{m} P_{\mathbb{X}|N}(k|n)\, P_N(n)$$

- $$= \sum_{n=k}^{m} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \frac{m!}{n!(m-n)!} q^n (1-q)^{m-n}$$

$\vdots$ (tedious computation)

$$= \frac{m!}{k!(m-k)!} (pq)^k (1-pq)^{m-k} \qquad k=0,\dots,m$$

So $\mathbb{X}$ has a binomial distribution with parameters pq, m.

We sometimes want to use the cumulative distribution function rather than the p.m.f.

Definition 1.3.4

Let $\mathbb{X}$ be an integer valued, non negative random variable. The distribution function of $\mathbb{X}$ is

$$F_{\mathbb{X}}(x) = F(x) = P(\mathbb{X} \le x) = \sum_{j \le x} p_j$$

Distribution functions have many properties. For example F is a piecewise constant, monotone increasing function, $F(x)=0, x\le 0$, F changes values at integers, $F(x)=1$ $x\to\infty$. We have various formulas such as

- $p_i = F(i) - \lim_{x\uparrow i} F(x)$, $\quad P(i < \mathbb{X} \le k) = F(k) - F(i)$

Following the ideas above,

Definition 1.3.5

Let $\mathbb{X}, \mathbb{Y}$ be integer valued, nonnegative random variables. The conditional distribution function of $\mathbb{X}$ given $\mathbb{Y}=k$ is

$$F_{\mathbb{X}|\mathbb{Y}}(x|k) = \frac{P(\mathbb{X} \le x, \mathbb{Y}=k)}{P(\mathbb{Y}=k)}$$

when $P(\mathbb{Y}=k) > 0$ and is undefined otherwise

Theorem 1.3.7

$F_{\mathbb{X}|\mathbb{Y}}(x|k)$ is a distribution function in x for each k.

Proof

Good exercise

Theorem 1.3.8

Let $\mathbb{X}, \mathbb{Y}$ be integer valued, nonnegative random variables.

(1.3.7) $$F_{\mathbb{X}|\mathbb{Y}}(x|k) = \sum_{j \le x} P_{\mathbb{X}|\mathbb{Y}}(j,k)$$

Proof

Exercise

Example 1.3.4

Phone calls arrive at a mail order company such that the number of calls in a minute has a Poisson distribution with mean 4. A given caller has a probability .5 of being female, independent of other callers. In a given minute, let X be the number of female callers and Y the total number of callers. Compute the joint p.m.f. of X and Y, the marginal p.m.f.s., and the conditional p.m.f of Y given $X = j$.

X and Y are not independent since $X \le Y$ always.

If $Y = k$, the number of female callers is binomial with parameters .5, k,

$$P_{X,Y}(j,k) = P(X = j \mid Y = k)\, P(Y = k)$$

be sure to understand

$$= \binom{k}{j}\left(\tfrac{1}{2}\right)^{j}\left(\tfrac{1}{2}\right)^{k-j} e^{-4}\frac{4^k}{k!}$$

$$= e^{-4}\frac{2^k}{j!\,(k-j)!}, \qquad 0 \le j \le k,\ k = 0,1,2\ldots$$

We know $P_Y(k) = e^{-4}\frac{4^k}{k!}$. We compute

$$P_X(j) = e^{-4} \sum_{k=j}^{\infty} \frac{2^k}{j!(k-j)!} = e^{-4} \frac{2^j}{j!} \sum_{k=j}^{\infty} \frac{2^{k-j}}{(k-j)!} = e^{-2} \frac{2^j}{j!}$$

the summands for $k<j$ are zero!

so $X \sim Poi(2)$.

Conditioned on $X = j$, the range of Y is $j, j+1, \ldots$

The conditional p.m.f. is

$$P_{Y|X}(k|j) = \frac{e^{-4} 2^k/(j!(k-j)!)}{e^{-2} 2^j/j!} = e^{-2} \frac{2^{k-j}}{(k-j)!}, \quad k = j, j+1, \ldots$$

Given j female callers, the total number of callers is j plus a number that is $Poi(2)$, which is just the number of male callers.

Recall from Theorem 1.3.5 that $P_{X|Y}(x, k)$ is a p.m.f. in x for each fixed k. We can compute statistics of this.

<u>Definition 1.3.6</u>

Let X, Y be integer valued, non negative random variables. The <u>conditional expected value</u> of X given $Y = k$ is

- (1.3.8) $E(X|k) = E(X|Y=k) = \sum_{j=0}^{\infty} j \, P_{X|Y}(j|k)$

provided $P_Y(k) > 0$, and is undefined otherwise.

- We can view $E(X|Y=y)$ as a function of y which equals $E(X|Y=k)$ when $y = k$.

<u>Definition 1.3.7</u>

We write $E(X|Y)$ for this function and we call it the <u>conditional expectation</u> of X given Y.

— #3 1/29/08

<u>Example 1.3.5</u>

Let X have a uniform distribution on $0, 1, 2, \ldots, n$ and given $X = j$, let Y have a uniform distribution on $0, 1, 2, \ldots, j$.

We have $Y | X = j \sim \text{unif}\{0, 1, 2, \ldots, j\}$, so

$$E(Y|X=j) = \frac{j}{2}$$

This means that

$$E(Y|X) = \frac{X}{2}$$