Math 425
Intro to Probability
Lecture 31

Kenneth Harris
kaharri@umich.edu

Department of Mathematics
University of Michigan

April 3, 2009

Joint Expectation

☞ Let X and Y be continuous (discrete) random variables X and Y with joint density (mass) $f_{X,Y}(x,y)$. Let Z be a random variable determined by X and Y,

$$Z = g(X,Y).$$

☞ The expected value of Z is given by

$$E[Z] = \int_{-\infty}^{\infty} z \cdot f_Z(z)\, dz$$

However, we can compute $E[Z]$ directly from the known joint density (mass) $f_{X,Y}(x,y)$ and the defining function $g(x,y)$.

Joint Expectation

☞

Theorem

Let X, Y and Z be random variables such that

$$Z = g(X,Y).$$

☞ *If X and Y are discrete with joint mass function $p_{X,Y}(x,y)$, then*

$$E[Z] = \sum_x \sum_y g(x,y) \cdot p_{X,Y}(x,y).$$

☞ *If X and Y are continuous with joint density function $f_{X,Y}(x,y)$, then*

$$E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot f_{X,Y}(x,y)\, dx\, dy$$

Example: Darts

Example. A dart thrown at a circular target of radius 1 strikes at a point (X,Y) in the target at random.

(i) What is the expected distance $\sqrt{X^2+Y^2}$ from the center?

(ii) What is the expected value of X^2+Y^2?

☞ The joint density of X and Y is

$$f_{X,Y}(x,y) = \pi^{-1} \qquad x^2+y^2 \le 1.$$

Example – continued

☞ It is best to convert to polar coordinate

$$r = \sqrt{x^2 + y^2} \qquad dx\,dy = r\,dr\,d\theta \qquad 0 < r < 1,\ 0 \le \theta < 2\pi.$$

(i) $E[\sqrt{X^2 + Y^2}]$.

$$\begin{aligned} E[\sqrt{X^2+Y^2}] &= \int\int_{x^2+y^2\le 1} \sqrt{x^2+y^2}\pi^{-1}\,dx\,dy \\ &= \int_0^{2\pi}\int_0^1 \pi^{-1} r^2\,dr\,d\theta \\ &= \frac{2}{3}. \end{aligned}$$

Example – continued

☞ It is best to convert to polar coordinate

$$r = \sqrt{x^2 + y^2} \qquad dx\,dy = r\,dr\,d\theta \qquad 0 < r < 1,\ 0 \le \theta < 2\pi.$$

(ii) $E[X^2 + Y^2]$.

$$\begin{aligned} E[X^2+Y^2] &= \int\int_{x^2+y^2\le 1} (x^2+y^2)\,\pi^{-1}\,dx\,dy \\ &= \int_0^{2\pi}\int_0^1 \pi^{-1} r^3\,dr\,d\theta \\ &= \frac{1}{2}. \end{aligned}$$

☞ Note that it NOT true that $E[Z^2]$ is $E[Z]^2$:

$$E[X^2+Y^2] \ne \left(E[\sqrt{X^2+Y^2}]\right)^2.$$

Expectation of Sums

Theorem

If X and Y are random variables whose expectations exist, then

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

If $X_1, \ldots, X_n$ are random variables whose expectations exist, then

$$E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i].$$

☞ Recall that for any random variable X whose expectation exists and c is a constant, then

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

So, given constants $c_1, \ldots, c_n$ and the conditions of the theorem

$$E[\sum_{i=1}^n c_i X_i] = \sum_{i=1}^n c_i E[X_i].$$

Proof

☞ Here is the proof when X and Y are continuous with joint density $f_{X,Y}(x,y)$.

$$\begin{aligned} E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\cdot f_{X,Y}(x,y)\,dx\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(x\cdot f_{X,Y}(x,y) + y\cdot f_{X,Y}(x,y)\right)dx\,dy \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x\cdot f_{X,Y}(x,y)\,dy\right)dx + \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} y\cdot f_{X,Y}(x,y)\,dx\right)dy \\ &= \int_{-\infty}^{\infty} x\cdot f_X(x)\,dx + \int_{-\infty}^{\infty} y\cdot f_Y(y)\,dy \\ &= E[X] + E[Y]. \end{aligned}$$

The generalization to sums of n random variables is by induction.

Example: Dice

Example. Add the scores of n fair die whose values are independent of each other. What is the expected value of the sum X of the scores.

☞ Let X_i count the score of the ith die. Then

$$X = \sum_{i=1}^{n} X_i$$

Since $E[X_i] = \frac{7}{2}$,

$$E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \frac{7n}{2}.$$

Example: Negative Binomial Variables

Example. Consider a Bernoulli trials process with success probability p. Let $X(k)$ count the number of trials until k successes. So, $X(k)$ is negative binomial random variable.

☞ Let X_i ($i = 1, \ldots, k$) be geometric random variables counting the number of trials between the $i-1$st and ith success. From Chapter 4: $E[X_i] = \frac{1}{p}$.

In Chapter 6 we showed $X(k) = X_1 + \ldots + X_k$. So,

$$E[X(k)] = \sum_{i=1}^{k} E[X_i] = k \cdot \frac{1}{p} = \frac{k}{p}.$$

This is what we found in Section 4.8.2.

Indicator variables

☞ A random variable I is an indicator for the event A if

$$I(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \in A^c \end{cases}$$

Key Property. Let I be an indicator variable for E. Then

$$\begin{aligned} E[I] &= 0 \cdot \mathbf{P}\{A^c\} + 1 \cdot \mathbf{P}\{A\} \\ &= \mathbf{P}\{A\}. \end{aligned}$$

Use of Indicator variables

Use of Indicators. Suppose X is the number of events that occur among some collection of events $A_1, \ldots, A_n$. Let $X_1, \ldots, X_n$ be indicator variables for these events.

Then $X = \sum_{k=1}^{n} X_k$, so

$$E[X] = E[\sum_{k=1}^{n} X_k] = \sum_{i=1}^{k} E[X_k] = \sum_{i=1}^{k} \mathbf{P}\{A_k\}$$

When to use. The method of indicator variables is useful when we need to find $E[X]$ for a counting variable X, especially when we can break it down into counting the occurrences of events in a collection $A_1, \ldots, A_n$, where the probability $\mathbf{P}\{A_k\}$ is easy to compute.

Example: Aces in Hand

Example. Let X count the number of aces in a 5-card poker hand. What is $E[X]$?

☞ Let X_i ($i = 1, \ldots, 5$) be the indicator variable for the ith card being an ace.

The probability of any card being an ace is $\frac{4}{52}$, so $E[X_i] = \frac{4}{52}$.

☞ Since $X = X_1 + \ldots + X_5$,

$$E[X] = 5 \cdot \frac{4}{52} = \frac{5}{13}.$$

Compare this to a direct computation

$$E[X] = \sum_{k=0}^{4} k \cdot \mathbf{P}\{X = k\} = \sum_{k=0}^{4} k \cdot \frac{\binom{4}{k}\binom{48}{5-k}}{\binom{52}{5}}.$$

Example: Gift Giving

Example. A common strategy for exchanging gifts is to put everyone's name in a hat, and let each draw a name of a person to give a gift. Given N people are exchanging gifts, what is the expected number of people who will draw their own name?

☞ Let G_i be an indicator variable that the ith person draws their own name:

$$G_i = \begin{cases} 1 & \text{if person } i \text{ draws their own name,} \\ 0 & \text{otherwise.} \end{cases}$$

Each person is equally likely to draw any name, so $\mathbf{P}\{G_i = 1\} = \frac{1}{N}$.

☞ Let G count the number of people who draw their own name.

Then $G = \sum_{i=1}^{N} G_i$, so

$$E[G] = \sum_{i=1}^{n} E[G_i] = N \cdot \frac{1}{N} = 1.$$

Thus, 1 out of N people can be expected to draw their own name.

Example: Binomial means

Example. In a Bernoulli trials process with n trials, each of the trials are independent and have the same probability of success p.

Let B count the number of successes, so that B is a binomial random variable.

☞ Let B_i be the indicator variable for success in the ith trial, so $E[B_i] = p$.

Since $B = \sum_{i=1}^{n} B_i$,

$$E[B] = \sum_{i=1}^{n} E[B_i] = n \cdot p.$$

This is exactly what we computed in Chapter 4 (section 4.6.1).

Example: Nonhomogeneous Bernoulli trials

Example. Suppose we have a complicated structure with n different elements that either work or fail independently of each other with different probabilities. What is the expected number of working elements?

☞ Let X_i be an indicator variable that the ith element works. So, for some $p_i \in [0, 1]$ we have $E[X_i] = p_i$. (This is the probability the ith element works.)

☞ Let X count the number of elements that work.

Then $X = \sum_{i=1}^{n} X_i$, so

$$E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i.$$

Example: Hypergeometric mean

Example. Suppose n balls are drawn at random from an urn without replacement. There are r red balls and b blue balls where $n < r$ and $n < b$. What is the expected number of red balls drawn.

☞ Let R denote the number of red balls drawn, and R_i ($i \leq n$) be the indicator variable for the event that the ith ball drawn is red.

The ith ball drawn is equally likely to be any of the $r + b$ balls, so

$$E[R_i] = \frac{r}{r+b}$$

Since $R = \sum_{i=1}^{n} R_i$, the expected number of red balls drawn is

$$E[R] = \sum_{i=1}^{n} E[R_i] = \frac{nr}{r+b}$$

This is exactly what we found in Chapter 4 (section 4.8.3).

Example: Screeing

Example

The concept of screening pooled samples originated during the Second World War to detect syphilis in U.S. soldiers. The strategy is used for testing a large number of samples (typically blood).

- Subgroups of at most 15 samples are chosen.
- Part of each specimen from each subgroup is pooled and tested.
- If a subgroup tests negative, then all of the individual samples in that subgroup are declared negative.
- If a subgroup tests positive, then each constituent sample of the subgroup is subsequently tested individually.

The use of pooling has been more extensively used recently in HIV testing of large specimen pools of low risk populations, such as in blood banks.

Example – continued

☞ 1000 soldiers are to be tested for syphillis. They will be broken into subgroups of 10. The samples in each subgroup will be tested, and if negative, the entire subgroup will be declared negative. However, if positive, each person in the group will be tested individually.

Suppose 1% of the soldier population has syphillis.

Question. How many tests can be expected to be performed using pooling, compared with testing each specimen individually?

☞ Testing every specimen individually requires 1000 tests.

Example – continued

☞ Let Y_i ($i = 1, \ldots, 100$) be random variables which count the number of tests administered to the ith subgroup.

If some sample is positive, then 11 tests are performed, and otherwise only 1 test is performed. So, for each i

$$E[X_i] = 1 \cdot 0.99^{10} + 11 \cdot (1 - 0.99^{10}) \approx 1.956.$$

☞ Since each random variable is identically distributed, they have the same expectation. The number of tests expected is

$$E[X_1 + \ldots + X_{100}] \approx 100 \cdot 1.956 = 195.6.$$

By testing every individual we would require 1000 tests.

Example: Coupon Collecting Problem

Example

Each packet of some ineffably dull and noxious product contains one of N different types of flashy coupon. Each packet is equally likely to contain any of N types. How many packets should you expect to purchase until you first possess all N types?

☞ Let T be the number of packets you must purchase until you first possess all N types.

Break the problem down into simpler components.

- Let T_i ($i = 1, 2, \ldots, N-1$) be the number of further packets you purchased after you got the $i-1$st new coupon until you got a new coupon (your ith).

For example, $T_1 = 1$, since any coupon you get will be new (you started with 0).

Example – continued

☞ T_i is a geometric random variable, where the probability of success (you collect a new coupon) is $p_i = \frac{N-i+1}{N}$ (there are $i-1$ out of N that are failures). So,

$$E[T_i] = \frac{1}{p_i} = \frac{N}{N-i+1}.$$

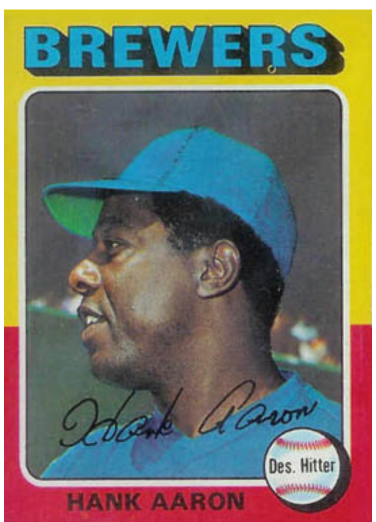
☞ Since $T = \sum_{i=1}^{N} T_i$, the expected number of packets is

$$\begin{aligned} E[T] &= \sum_{i=1}^{N} E[T_i] = \sum_{i=1}^{N} \frac{N}{N-i+1} \\ &= N\Big(\sum_{i=1}^{N} \frac{1}{i}\Big) \\ &\approx N\log N + N\gamma \qquad \text{where } \gamma \approx 0.5772 \end{aligned}$$

γ is known as the Euler-Masheroni constant. (It is unknown whether it is rational or not.)

Example – continued

☞ As a kid I desparately needed Topps #660, Hank Aaron, to complete my collection:



Example – continued

☞ There were 660 cards in the complete set of 1975 Topps Baseball Series. I think there were 15 cards (and one ineffably dull and noxious product) in a packet, which was about 15 cents.

I should expect to have to collect alot of cards before I collected each card:

$$E[T] \approx 660 \ln 660 + (0.5772)660 \approx 4666 \text{ cards}.$$

☞ \$46.66 is not chump change to a kid on a quarter a week allowance!!