Statistics 110: Introduction to Probability — Section 4, Fall 2013

Expected Value and Indicator Random Variables

William Chen (wchen@college.harvard.edu) — *Sebastian Chiu (schiu@college.harvard.edu)*

Key Topics

Expected Value. Very common thing to find in Statistics, as this weighted average is the one number that can (usually) best represent a whole distribution. Use **Linearity of Expectation** to find more complicated expectations, and remember that any two random variables that have the same distribution must have the same expectation by **symmetry**.

Indicator Random Variables. These simplify any counting problem and many discrete problems, as you can use indicator random variables to indicate the happening of any event. The **Fundamental Bridge** says that the expected value of an indicator random variable is always the probability that it happens.

More Practice. Check out Strategic Practice 4.

Distributions

Don't confuse a random variable with its distribution! A distribution is like a blueprint for a house, and the random variable is the house itself. Distributions can be described in a variety of ways, including the CDF or the PMF. The **support** of a random variable is the set of possible values that it can take.

Probability Mass Function (PMF) (Discrete Only) is a function that takes in the value x, and gives the probability that a random variable takes on the value x. The PMF is a positive-valued function, and $\sum_x P(X = x) = 1$

$$P_X(x) = P(X = x)$$

Cumulative Distribution Function (CDF) is a function that takes in the value x, and gives the probability that a random variable takes on the value at most x.

$$F(x) = P(X \leq x)$$

Expected Value, Linearity, and Symmetry

Expected Value (aka *mean*, *expectation*, or *average*) can be thought of as the "weighted average" of the possible outcomes of our random variable. Mathematically, if $x_1, x_2, x_3, \dots$ are all of the possible values that X can take, the expected value of X can be calculated as follows:

$$E(X) = \sum_i x_i P(X = x_i)$$

Note that for *any* X and Y, a and b scaling coefficients and c is our constant, the following property of **Linearity of Expectation** holds:

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

If two Random Variables have the same distribution, *even when they are dependent* by the property of **Symmetry** their expected values are equal.

Conditional Expected Value is calculated like expectation, only conditioned on any event A.

$$E(X|A) = \sum_x xP(X = x|A)$$

Indicator Random Variables

Indicator Random Variables is random variable that takes on either 1 or 0. The indicator is always an indicator of some event. If the event occurs, the indicator is 1, otherwise it is 0. They are useful for many problems that involve counting and expected value.

Distribution $I_A \sim \text{Bern}(p)$ where $p = P(A)$

Fundamental Bridge The expectation of an indicator for A is the probability of the event. $E(I_A) = P(A)$. Notation:

$$I_A = \begin{cases} 1 & \text{A occurs} \\ 0 & \text{A does not occur} \end{cases}$$

Bernoulli and Binomial Distributions

Bernoulli The Bernoulli distribution is the simplest case of the Binomial distribution, where we only have one trial, or $n = 1$. Let us say that X is distributed Bern(p). We know the following:

Story. X "succeeds" (is 1) with probability p, and X "fails" (is 0) with probability $1 - p$.

Example. A fair coin flip is distributed Bern($\frac{1}{2}$).

PMF. The probability mass function of a Bernoulli is:

$$P(X = x) = p^x(1 - p)^{1-x}$$

or simply

$$P(X = x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \end{cases}$$

Binomial Let us say that X is distributed Bin(n, p). We know the following:

Story X is the number of "successes" that we will achieve in n independent trials, where each trial can be either a success or a failure, each with the same probability p of success. We can also say that X is a sum of multiple independent $Bern(p)$ random variables. Let $X \sim \text{Bin}(n, p)$ and $X_j \sim \text{Bern}(p)$, where all of the Bernoullis are independent. We can express the following:

$$X = X_1 + X_2 + X_3 + \cdots + X_n$$

Example If Jeremy Lin makes 10 free throws and each one independently has a $\frac{3}{4}$ chance of getting in, then the number of free throws he makes is distributed Bin($10, \frac{3}{4}$), or, letting X be the number of free throws that he makes, X is a Binomial Random Variable distributed Bin($10, \frac{3}{4}$).

PMF The probability mass function of a Binomial is:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

Binomial Coefficient $\binom{n}{k}$ is a function of n and k and is read n *choose* k, and means out of n possible distinguishable objects, how many ways can I possibly choose k of them? The formula for the binomial coefficient is:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Geometric, Negative Binomial, and Hypergeometric Distributions

Geometric Let us say that X is distributed Geom(p). We know the following:

- **Story** X is the number of "failures" that we will achieve before we achieve our first success. Our successes have probability p.
- **Example** If each pokeball we throw has a $\frac{1}{10}$ probability to catch Mew, the number of failed pokeballs will be distributed Geom($\frac{1}{10}$).
- **PMF** With $q = 1 - p$, the probability mass function of a Geometric is:

$$P(X = k) = q^k p$$

Negative Binomial Let us say that X is distributed NBin(r, p). We know the following:

- **Story** X is the number of "failures" that we will achieve before we achieve our rth success. Our successes have probability p.
- **Example** Thundershock has 60% accuracy and can faint a wild Raticate in 3 hits. The number of misses before Pikachu faints Raticate with Thundershock is distributed NBin($3, .6$).
- **PMF** With $q = 1 - p$, the probability mass function of a Negative Binomial is:

$$P(X = n) = \binom{n + r - 1}{r - 1} p^r q^n$$

Hypergeometric Let us say that X is distributed Hypergeometric(w, b, n). We know the following:

- **Story** In a population of b undesired objects and w desired objects, X is the number of "successes" we will have in a draw of n objects, without replacement.
- **Example** 1) Let's say that we have only b Weedles (failure) and w Pikachus (success) in Viridian Forest. We encounter n of the Pokemon in the forest, and X is the number of Pikachus in our encounters. 2) The number of aces that you draw in 5 cards (without replacement). 3) You have w white balls and b black balls, and you draw b balls. X is the number of white balls you will draw in your sample. 4) Elk Problem - You have N elk, you capture n of them, tag them, and release them. Then you recollect a new sample of size m. How many tagged elk are now in the new sample?
- **PMF** The probability mass function of a Hypergeometric is:

$$P(X = k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$$

Discrete Distributions

Distribution	PDF and Support	Expected Value	Equivalent To
Bernoulli Bern(p)	$P(X = 1) = p$ $P(X = 0) = q$	p	Bin($1, p$)
Binomial Bin(n, p)	$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ $k \in \{0, 1, 2, \dots n\}$	np	Sum of n independent Bern(p)
Geometric Geom(p)	$P(X = k) = q^k p$ $k \in \{0, 1, 2, \dots\}$	$\frac{q}{p}$	NBin($1, p$)
Negative Binomial NBin(r, p)	$P(X = n) = \binom{n+r-1}{r-1} p^r q^n$ $n \in \{0, 1, 2, \dots\}$	$r\frac{q}{p}$	Sum of r Geom(p)
Hypergeometric Hypergeometric(w, b, n)	$P(X = k) = \binom{w}{k}\binom{b}{n-k}/\binom{w+b}{n}$ $k \in \{0, 1, 2, \dots\}$	$n\frac{w}{b+w}$	

Practice Problems

Example 1. A Betting Game We play a game where you give me a number and then I roll two standard dice. If the sum of the two die equals the number you give me, you win the amount of your number *squared* in dollars. What number should you choose (to maximize expected return)?

Example 2. Gotta Catch'em All Your pokeball has a constant p chance of success against Mew. You plan to keep throwing pokeballs until you catch it. What is the expected number of pokeballs that will fail before you catch Mew?

Example 3. College Application Mishap A lazy high school senior types up applications and envelopes to n different colleges. The letters are randomly put into the envelopes.

a) On average, how many applications went to the right college?

b) A match occurs if college A receives college B's envelope AND if college B receive college A's envelope. On average, how many matches occur?

Example 4. First Day of School It's the first day of class and 15 students who don't know each other seat themselves randomly around a table. 8 of those students are Riverlings while the other 7 are Quadlings. If two students seated adjacently live in the same area, they will become friends. On average, how many pairs of friends do you expect to be formed at the table?

Example 5. Maximizing Profits You pay \$1 per turn to roll a fair 100-sided die repeatedly. When you decide to stop rolling, you get an amount in dollars equal to the face value of the die. What is your optimal strategy? What are your expected earnings under the optimal strategy?