STATS 116 Review *Notation and Definitions.* We briefly recall some facts and notation from STATS 116 that will be useful in reading the solutions below.

Expected value. Let X be a random variable. Suppose X is **discrete**, that is, the values that X can take on can be enumerated in a list (possibly of infinite length) as $x_1, x_2, \ldots$. Then the expected value of X is

$$E(X) = \sum_{n=1}^{\infty} x_n P(X = x_n) .$$

If X is **continuous**—meaning, roughly, that the possible values it takes on cannot be enumerated in a list—with density function $f_X(x)$, then the expected value is defined by the expression

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx .$$

These are the only two classes of random variables we will encounter in this class.

Independence. Two random variables X and Y are called (pairwise) **independent** if

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$$

for all choices of x and y. The intuitive notion behind independence is that knowing the value of X tells you nothing about the value of Y, and vice versa.

For discrete random variables, an equivalent (and usually easier to verify) condition for independence is that

$$P(X = x_n, Y = y_m) = P(X = x_n)P(Y = y_m) \text{ for all } x_n,\, y_m.$$

For continuous random variables, an equivalent condition is that the joint density function $f_{X,Y}(x, y)$ factors,

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) .$$

Straightforward calculus shows that if X and Y are independent, then $E(XY) = E(X)E(Y)$. However, the converse is ***not true***! That is, in general, just because $E(XY) = E(X)E(Y)$ for some pair of random variables does not mean that X and Y are independent.

Independence can be extended beyond pairs of random variables. The set of random variables $X_1, X_2, \ldots, X_n$ are (jointly) independent if and only if

$$P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) = P(X_1 \le x_1)P(X_2 \le x_2)\cdots P(X_n \le x_n) .$$

The factorization criteria for discrete and continuous random variables generalize in the obvious way. It is important to note that pairwise independence of all pairs X_i, X_j, $i \ne j$, does *not*, by itself, imply joint independence.

Bernoulli random variable. A Bernoulli random variable with probability of success p, denoted by Ber(p) is a random variable taking on the two values 0 and 1, and is defined as

$$X = \begin{cases} 0, & \text{with probability } 1 - p \\ 1, & \text{with probability } p . \end{cases}$$

Bernoulli random variables are discrete. Using the definition given above, the expected value of a Ber(p) random variable is

$$E(X) = 0 \cdot (1 - p) + 1 \cdot p = p .$$

The outcome of a single fair coin flip is modeled as a Ber(1/2) random variable.

Problem 1 *Integral.* Evaluate $\int_0^1 e^{at}\,dt$.

Solution.

$$A = \int_0^1 e^{at}\,dt = a^{-1}e^{at}\Big|_0^1 = \frac{1}{a}(e^a - 1)$$

Problem 2 *Expected value.* Which of the following statements are always true?

(a) $E(XY) = E(X)E(Y)$

Solution. False. Pairs of random variables that satisfy the above relation are called **uncorrelated**. Random variables that are pairwise **independent** are a special case of uncorrelated variables.

Here is a simple example of failure of the equation above. Let $X \sim \text{Ber}(1/2)$, and $Y = X + 1$. Then,

$$\begin{aligned} E(X) &= 0 \cdot (1/2) + 1 \cdot (1/2) = 1/2\,, \\ E(Y) &= E(X+1) = 1 \cdot (1/2) + 2 \cdot (1/2) = 3/2\,, \\ E(XY) &= E(X(X+1)) = E(X^2 + X) = 0 \cdot (1/2) + 2 \cdot (1/2) = 1\,, \end{aligned}$$

from which we clearly see that $1 = E(XY) \neq E(X)E(Y) = 3/4$.

(b) $E(X+Y) = E(X) + E(Y)$

Solution. True (almost). Actually, we need a *very* minor technical condition: at least one of $E(X)$ and $E(Y)$ must be finite. This condition is true for (virtually) any pair of random variables normally encountered in practice, and will *always* be true for the random variables encountered in this class. Thus, for the purposes of this class—and likely anything you will ever do with probability—the answer is "true". (If you are *really* interested in an example where this doesn't hold, you can e-mail the TA.)

A very slight generalization of the above expression is that, for any constant a,

$$E(aX + Y) = aE(X) + E(Y)\,.$$

This important property of expected values is called **linearity**.

Here is a quick proof of the above for the special case of a joint distribution function with a density. Let $f(x, y)$ be the joint probability density function of (X, Y), $f_X(x)$ be the marginal density of X and $f_Y(y)$ be the marginal density of Y. Clearly,

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\,dy\,, \text{ and,} \\ f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx\,. \end{aligned}$$

Then

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

where the last line follows from the definition of $E(X)$ and $E(Y)$.

(c) $E\frac{X}{Y} = \frac{EX}{EY}$.

Solution. False. Here is a counterexample. Let $X = 1$ and let

$$Y = \begin{cases} 1, & \text{with probability } 1/2 \\ 2, & \text{with probability } 1/2 \end{cases} .$$

Then $E(X/Y) = E(1/Y) = 1 \cdot (1/2) + (1/2) \cdot (1/2) = 3/4$. But, $\frac{EX}{EY} = \frac{1}{EY} = \frac{1}{3/2} = 2/3$.

Indeed, for the case where X and Y are independent, $E(X) \neq 0$ and Y is strictly positive (or strictly negative), the above statement is *never* true. This is a consequence of what is known as **Jensen's inequality**.

However, equality can occur, for example, if we force $E(X) = 0$. Below is an example. Let

$$X = \begin{cases} 1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \,. \end{cases}$$

and

$$Y = \begin{cases} 1, & \text{with probability } 1/2 \\ 2, & \text{with probability } 1/2 \,. \end{cases}$$

with X and Y being independent. Clearly $E(X) = 0$ and $E(Y) = 3/2$. Also

$$E\frac{X}{Y} = \frac{1}{4}\left(\frac{-1}{1} + \frac{1}{1} + \frac{-1}{2} + \frac{1}{2}\right) = 0 \,.$$

So, we have that $E\frac{X}{Y} = \frac{E(X)}{E(Y)} = 0$.

Problem 3 *Waiting time.* What is the expected number of fair coin tosses required to get the first head? For example if the coin tosses are TTHTH..., then the number of tosses is 3.

Solution. Intuitively, we may guess that if there are n equally likely outcomes and we perform independent trials, then it should take about n trials before we see any particular outcome, on average. Hence, we might guess that in the coin-flipping case, the answer should be 2.

We now verify this in three different ways. The first method is purely computational, the second is more intuitive, and the third uses a standard result from STATS 116. You don't need to understand all three, but you should try to understand at least one of them.

Method 1: Direct method. The probability of a head occurring on the nth trial is $(1 - 1/2)^{n-1}(1/2) = 2^{-n}$, since trials are independent and the first $n - 1$ trials must be tails and the last one must be a head. If X is the random variable corresponding to the number of trials, then X is discrete and takes on values 1, 2, ..., so

$$E(X) = \sum_{n=1}^{\infty} n2^{-n}.$$

But

$$\sum_{n=1}^{\infty} n2^{-n} - \sum_{n=1}^{\infty}(n-1)2^{-n} = \sum_{n=1}^{\infty} 2^{-n} = 1 \,,$$

where the last equality follows from the geometric series, i.e., for $|r| > 1$, $\sum_{n=1}^{\infty} r^{-n} = 1/(r-1)$.

On the other hand,

$$\sum_{n=1}^{\infty} n2^{-n} - \sum_{n=1}^{\infty}(n-1)2^{-n} = \sum_{n=1}^{\infty} n2^{-n} - \frac{1}{2}\sum_{n=1}^{\infty}(n-1)2^{-(n-1)} = \sum_{n=1}^{\infty} n2^{-n} - \frac{1}{2}\sum_{n=1}^{\infty} n2^{-n} = \frac{1}{2}E(X) \,.$$

Hence, $E(X) = 2$.

Method 2: First-step analysis. With probability 1/2, we will see heads on the very first flip, i.e., with X defined as before, we have that $X = 1$ with probability 1/2. Also, with probability 1/2, it will take more than one flip. In this case, we see that it will take $1 + Y$ flips, where Y is a random variable corresponding to the number of trials until the first heads. That is Y has the same distribution as X. Thus,

$$E(X) = 1 \cdot \frac{1}{2} + \frac{1}{2}\left(1 + E(Y)\right) ,$$

by linearity of expectation. Since X and Y have the same distribution, then $E(Y) = E(X)$, from which we get the recurrence

$$E(X) = \frac{1}{2}(2 + E(X)) ,$$

i.e., $E(X) = 2$.

This approach can be easily extended to biased coin flips, which is a more general case of having n equally likely events. Suppose that a biased coin lands heads with probability p and tails with probability $1 - p$. See if you can write out the steps to show that $E(X) = 1/p$.

Method 3: Moment-generating functions. If you have taken STATS 116 (or similar), you are familiar with the **moment-generating function** $m(t)$ of a random variable X. For our X we have

$$m(t) = E(e^{tX}) = \sum_{n=1}^{\infty} e^{tn} 2^{-n} = \sum_{n=1}^{\infty} (2e^{-t})^{-n} = \frac{e^t}{2 - e^t} ,$$

where, again, the last equality is from the geometric-series result.

Then, by the moment-generating function property, $E(X) = m'(0)$. Since

$$m'(t) = \frac{2e^t}{(2 - e^t)^2} ,$$

we get that $E(X) = m'(0) = 2$.

Problem 4 *Conditional Expectation.* Let X and Y be independent and identically distributed. Find $E(X|X+Y)$.

Solution. The trick to solving this problem is to add and subtract a quantity inside the conditional expectation to account for the fact that you are conditioning on $X + Y$. To this end we can write $E(X + Y|X + Y)$ in two ways,

$$E(X + Y|X + Y) = X + Y,$$

and

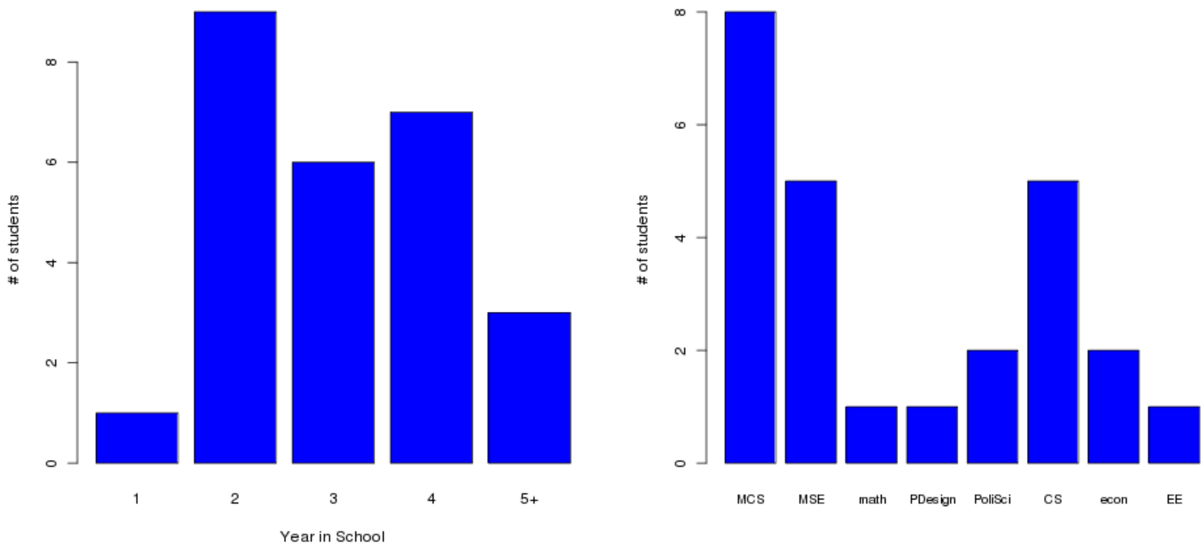
$$E(X + Y|X + Y) = E(X|X + Y) + E(Y|X + Y) = E(X|X + Y) + E(X|X + Y) = 2E(X|X + Y),$$

where we used the fact that the conditional expectation of a variable, given the variable, is just itself for the first equation. For the second equation, we use linearity of the conditional expectation and that X and Y are identically distributed, and hence have the same conditional expectation. Equating these two results and dividing by 2 gives

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

This result should make sense intuitively. The best guess for the expectation of a random variable given the sum of two draws is the average of those draws.

Problem 5 *Questionnaire Statistics.*



(a) (b)

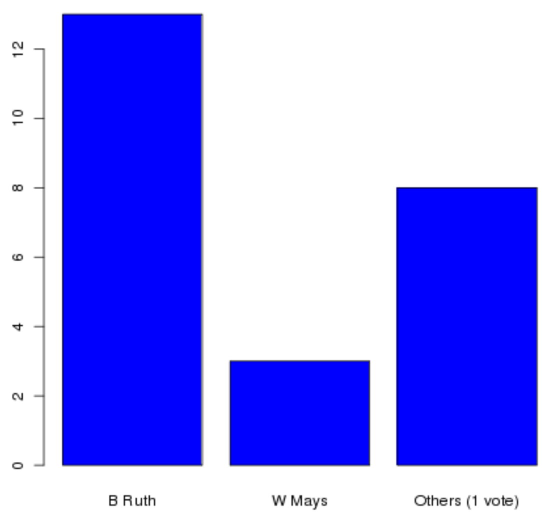
Figure 1: Distribution of Majors and Year in School of Students in Stats 50, 2011

Varsity College	Club College	HS Athlete	International / Olympic
4 (15)	3 (12)	20 (77)	1 (4)

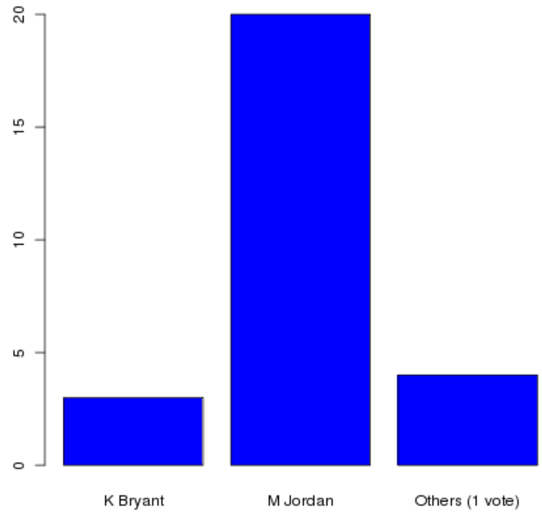
Table 1: Number (percentage) of self-reported atheltes at different levels. (If you didn't specify which level, I didn't count you. Sorry!)

Q1	Q2	Q3	Q4
19 (73)	20 (77)	19 (73)	3 (12)

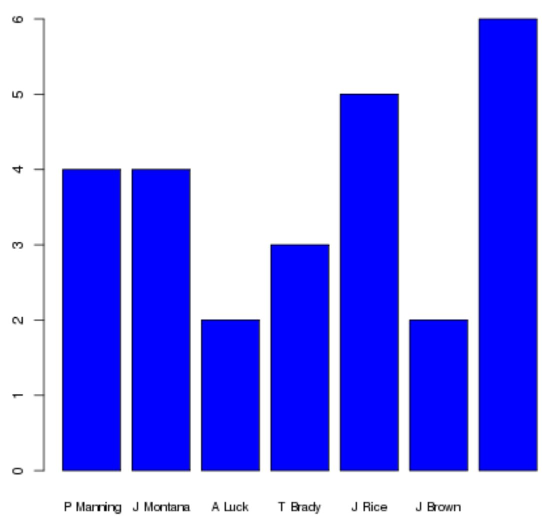
Table 2: Number (percentage) getting diagnostic question right. (26 total responses)



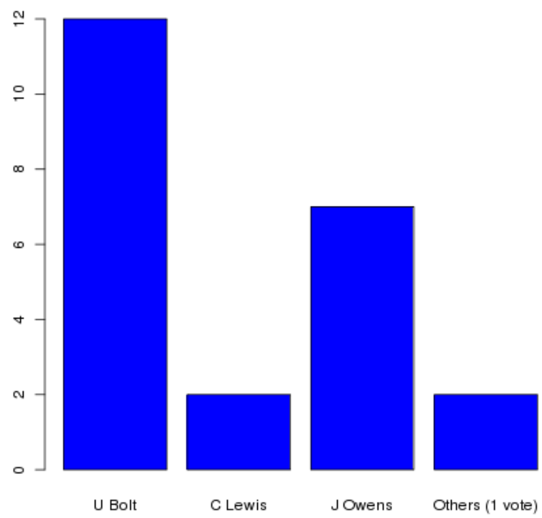
(a)



(b)



(c)



(d)

Figure 2: Votes for greatest player. Other players for Baseball: Lincecum, Gehrig, Griffey Jr., Paige, Williams, Sosa— Basketball: Abdul Jabbar, Olajuwon, Duncan, James— Football: Unitas, Elway, Payton, Sandars, Vick