

MIT OpenCourseWare
<http://ocw.mit.edu>

14.30 Introduction to Statistical Methods in Economics
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

Problem Set #2

14.30 - Intro. to Statistical Methods in Economics

Instructor: Konrad Menzel

Due: Tuesday, February 24, 2009

Question 1

Recall that a random variable X has the binomial distribution if

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

where n is the number of trials and p is the chance of success. For the following crudely-disguised questions about the binomial distribution, do the following: Define p , n , what the specific “trials” are, what “success” is. Then write down the relevant distribution and answer the specific question.

1. If 25 percent of the balls in a certain box are red, and if 15 balls are selected from the box at random, with replacement, what is the probability that more than four red balls will be obtained?
2. (3 pts.) Suppose an economist is organizing a survey of American minimum wage workers, and is interested in understanding how many workers that earn the minimum wage are teenagers.¹ Suppose further that one out of every four minimum wage workers is a teenager. If the economist finds 80 minimum wage workers for his survey, what’s the probability that he interviews exactly 14 teenagers? 35 teenagers? What’s the probability that he gets at least 5 teenagers in his survey?
3. (Bonus Question) A city has 5000 children, including 800 who have not been vaccinated for measles. Sixty-five of the city’s children are enrolled in a day care center. Suppose the municipal health department sends a doctor and nurse to the day care center to immunize any child who has not already been vaccinated. Find a formula for the probability that exactly k of the children at the day care center have not been vaccinated. (Hint: This is not exactly a binomial distribution problem.)

¹In the debate over the minimum wage, one point that is always made is that a lot of workers on the minimum wage are middle-class teenagers. Since most of them are not supporting a family, the harms of an increase in the minimum wage outweigh the benefits. This point never convinces anyone.

Question Two

Suppose you flip a weighted coin (probability of heads is p and probability of tails is $q = 1 - p$) n times.

1. What is the probability that you get a particular ordering of k heads and $n - k$ tails?
2. What is the probability that you get k heads and $n - k$ tails?
3. Let X be the number of heads in n flips. What is the probability density function of X ?
4. What does this question have to do with the MIT Beaver's football team question from Problem Set #1? Explain.

Suppose now that you have a hat with two coins, one weighted as above and one fair. You choose one at random and flip that one n times.

1. Let Y be the number of heads in n flips. What is the PDF of Y ?
2. What is the probability that you chose the fair coin given that $Y = k$?

Question Three

Suppose that two balanced dice are rolled. Determine and sketch the probability distribution of each of the following random variables.

1. Let X denote the absolute value of the difference between the two numbers that appear.
2. Let Y denote the product of the two numbers that appear.
3. Let Z denote the number of even numbers that appear.

Question Four

Suppose that you have just purchased a new battery for your smoke detector, and the life of the battery is a random variable X , with pdf

$$f_X(x) = ke^{-x/\beta},$$

where $x \in (0, \infty)$. Assume that t and s are real non-negative numbers.

1. Use the properties of a pdf to find the value of k .
2. Find an expression for $P(X \geq t)$.

3. Find an expression for $P(X \geq t + s | X \geq s)$.
4. Suppose that your batteries have lasted s weeks without dying. Based on your above answers, are you more concerned that the battery is about to die that you were when you first put it in?

Question Five

Suppose we investigate the pattern of genetic inheritance for the color of an exotic flower which has either blue or red blossoms. Since the flower lives in a close symbiotic relationship with the very shy squirrel monkey which can't be held in captivity, there is no way of doing a controlled laboratory experiment to answer the research questions.

Each flower carries the color genes of both its "father" and its "mother," so its genetic information can be described as a pair of genes (G_M, G_F) as given by the following table:

		"Father"	
		B	R
"Mother"	B	(B,B)	(B,R)
	R	(R,B)	(R,R)

The phenotype corresponding to red blossoms, R, is said to be *dominant* if any flower which contains at least one gene of the R type (e.g. the combination (R,B)) has red blossoms. Either the blue or the red phenotype is dominant, but before having seen a single specimen of the flower, we think that each possibility is equally likely.

1. Suppose we know beforehand that the R and the B alleles are equally frequent, i.e. $P_i(B) = P_i(R) = \frac{1}{2}$ for $i = F, M$, and independent across the "parents," i.e. $P_{FM}(G_F, G_M) = P_F(G_F)P_M(G_M)$. If the blue phenotype B is dominant, what is the probability that a given specimen of the flower has red blossoms? What is the probability of red blossoms if the R is dominant?
2. It takes a lot of effort to find a single specimen of the plant, so all a well-funded two-month expedition by a team of MIT botanists to the Amazon could gather was a sample of 15 flowers. If R is dominant, what is the probability of 9 out of the 15 flowers having red blossoms?
3. Our expedition did in fact return with a sample of 9 red and 6 blue blossoms. Given that, what is the likelihood that the red phenotype is dominant?
4. At the same time, there is a lonely graduate student working in the same area for two entire years on the same research question. The graduate student is totally cut off from the outside world and doesn't know about the other expedition's findings yet, but bases his inference solely on his own sample. What is his posterior probability of R being dominant given that he found N flowers, out of which x have red blossoms? Show that this probability does not depend directly on N , but only the difference between the number x of red blossoms, and the number $N - x$ of blue blossoms.