

## Problem Set 3

Reading: Random variables, mean, and LOTUS

Quiz Date: Tuesday, June 19

**Problem 1**

Suppose that in an experiment three fair coins are tossed. Let  $Y$  denote the number of heads appearing. Find the probability mass function (pmf) of  $Y$  and  $E[Y]$ .

**Solution**

$Y$  takes on one of the values 0, 1, 2, and 3.

$$\begin{aligned} p_Y(0) &= P(Y=0) = P(\{(T, T, T)\}) = 1/8, \\ p_Y(1) &= P(Y=1) = P(\{(T, T, H), (T, H, T), (H, T, T)\}) = 3/8, \\ p_Y(2) &= P(Y=2) = P(\{(T, H, H), (H, T, H), (T, H, H)\}) = 3/8, \\ p_Y(3) &= P(Y=3) = P(\{(H, H, H)\}) = 1/8. \end{aligned}$$

*Check:* The sum of probabilities should add up to one:  $1/8 + 3/8 + 3/8 + 1/8 = 8/8 = 1$ .

Before finding  $E[Y]$ , try to guess what its value!

For  $E[Y]$ , we have

$$\begin{aligned} E[Y] &= 0 \cdot p_Y(0) + 1 \cdot p_Y(1) + 2 \cdot p_Y(2) + 3 \cdot p_Y(3) \\ &= 0 \cdot 1/8 + 1 \cdot 3/8 + 2 \cdot 3/8 + 3 \cdot 1/8 \\ &= 12/8 \\ &= 3/2. \end{aligned}$$

**Problem 2**

We have a bag with 20 balls, numbered from 1 through 20. We pick a ball at random, observe its number and put it back. We do the same operation two more times. (This is called *sampling with replacement* since we replace the object that we pick – it may be chosen again.) In the end, we have three numbers. Let  $X$  denote the largest of these. Find the pmf of  $X$ . What is the probability that  $X \geq 17$ ?

**Solution**

There are 20 options for each of the three numbers that we observe. So  $|\Omega| = 20^3$ . Let us count the number of outcomes whose largest number is  $i$ . There are three cases to consider.

- Only one of the three numbers is  $i$ , and the other two are less than  $i$ : we have 3 ways to choose which number (that is, the first, the second, or the third number) is  $i$ . For the remaining numbers, we have  $(i-1)^2$  choices. So there are  $3(i-1)^2$  such cases.
- Two of the three numbers are  $i$ , and the other is less than  $i$ : we have 3 ways to choose which number is less than  $i$  and then we have  $i-1$  options for the value of this number. So there are  $3(i-1)$  such cases.
- All three numbers are equal to  $i$ . There is one such case.

Hence  $|\{X = i\}| = 3(i-1)^2 + 3(i-1) + 1 = 3i(i-1) + 1$  and so

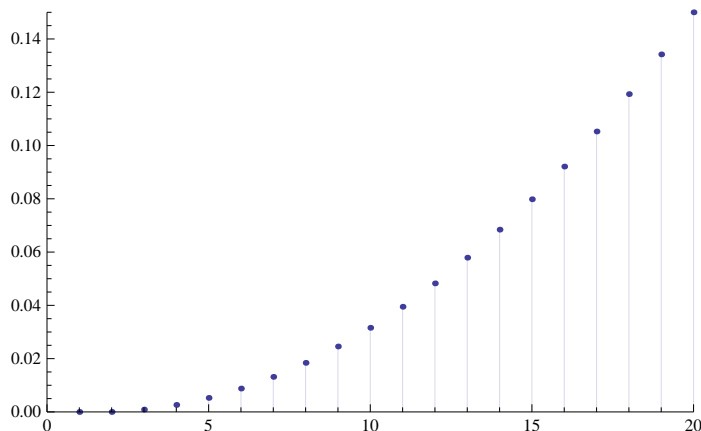
$$p_X(i) = \frac{3i(i-1) + 1}{20^3} \quad \text{for } i = 1, 2, \dots, 20.$$

Another way to see this: there are  $i^3$  triples of numbers such that no number exceeds  $i$ . There are  $(i-1)^3$  triples of numbers such that no number exceeds  $i-1$ . So there are  $i^3 - (i-1)^3$  triples of numbers such that the largest number is equal to  $i$ . Hence

$$p_X(i) = \frac{i^3 - (i-1)^3}{20^3} \quad \text{for } i = 1, 2, \dots, 20.$$

It can be easily checked that both answers are the same.

The plot of the pmf is below.



For the second part,

$$P(X \geq 17) = p_X(17) + p_X(18) + p_X(19) + p_X(20) = 0.488.$$

### Problem 3

We have a bag with 20 balls, numbered from 1 through 20. We pick three balls at random and observe their numbers. (This is called *sampling without replacement* since we do not replace the objects that we pick.) Let  $Y$  denote the largest of the numbers on the chosen balls. Find the pmf of  $Y$ . What is the probability that  $Y \geq 17$ ?

### Solution

In many cases, there are two ways to solve problems of sampling without replacement. The answer is the same from both methods:

#### 1) Order does not matter

This is generally the easier method.

There are  $\binom{20}{3}$  ways to choose three objects out of a set of size 20. So  $|\Omega| = \binom{20}{3}$ . Let us count the number of outcomes where the largest number is  $i$ . We need only consider one case: one of the numbers is  $i$  and the other two are smaller than  $i$ . There are  $\binom{i-1}{2}$  ways to choose two numbers smaller than  $i$ . Hence,

$$p_Y(i) = \frac{\binom{i-1}{2}}{\binom{20}{3}}, \quad \text{for } i = 1, 2, \dots, 20.$$

## 2) Order matters

The method here is similar to the method of the previous problem.

Suppose we consider the order to be important. Note that this is not to say that we are considering another experiment – this is just another way of looking at the same experiment.

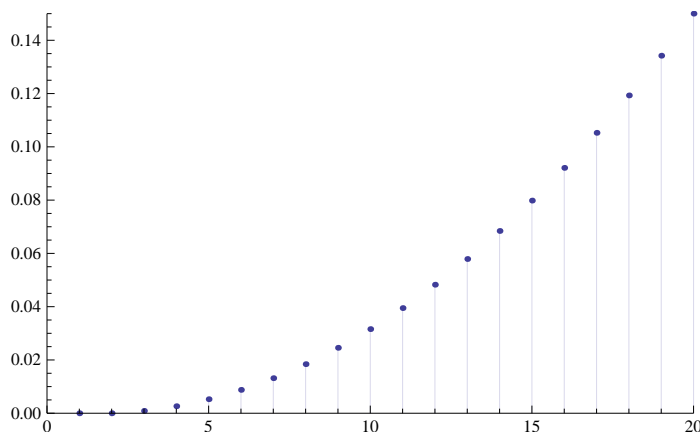
There are 20 ways to choose the first ball, 19 to choose the second and 18 to choose the third. So  $|\Omega| = 20 \cdot 19 \cdot 18$ . Let us count the number of the ways that the largest number is  $i$ . We have three ways to choose which number is  $i$  and we have  $(i-1)(i-2)$  choices for the remaining numbers since numbers cannot be repeated. So

$$p_Y(i) = \frac{3(i-1)(i-2)}{20 \cdot 19 \cdot 18}$$

Note that the solution is the same since

$$\frac{3(i-1)(i-2)}{20 \cdot 19 \cdot 18} = \frac{\frac{(i-1)(i-2)}{2}}{\frac{20 \cdot 19 \cdot 18}{3 \cdot 2}} = \frac{\binom{i-1}{2}}{\binom{20}{3}}.$$

The plot of the pmf is below.



For the second part,

$$P(Y \geq 17) = p_Y(17) + p_Y(18) + p_Y(19) + p_Y(20) = 0.509.$$

It is noteworthy that the probabilities are similar to those of the previous problem. However, here the probability that the maximum is large is slightly higher. Why?

## Problem 4

Three balls are randomly chosen from an urn containing 3 white, 3 red, and 5 black balls. Suppose that we win \$1 for each white ball selected and lose \$1 for each red ball selected. Let  $X$  be the total winning of the experiment. Find the pmf of  $X$  and  $E[X]$ .

## Solution

The possible values are  $0, \pm 1, \pm 2, \pm 3$ . It may be helpful to imagine that the balls are numbered:

$$B_1, B_2, B_3, B_4, B_5, W_1, W_2, W_3, R_1, R_2, R_3.$$

For  $i \in \{0, 1, -1, 2, -2, 3, -3\}$ ,

$$p_X(0) = \frac{\binom{5}{3} + \binom{5}{1} \binom{3}{1} \binom{3}{1}}{\binom{11}{3}} = \frac{55}{165},$$

$$p_X(1) = p_X(-1) = \frac{\binom{5}{2} \binom{3}{1} + \binom{3}{2} \binom{3}{1}}{\binom{11}{3}} = \frac{39}{165},$$

$$p_X(2) = p_X(-2) = \frac{\binom{5}{1} \binom{3}{2}}{\binom{11}{3}} = \frac{15}{165},$$

$$p_X(3) = p_X(-3) = \frac{\binom{3}{3}}{\binom{11}{3}} = \frac{1}{165}.$$

*Check:*  $\frac{55}{165} + \frac{39}{165} + \frac{39}{165} + \frac{15}{165} + \frac{15}{165} + \frac{1}{165} + \frac{1}{165} = 1.$

By symmetry, we expect the mean to be zero. The following calculations verify that.

$$E[X] = 0 \cdot \frac{55}{165} + (1) \frac{39}{165} + (-1) \frac{39}{165} + (2) \frac{15}{165} + (-2) \frac{15}{165} + (3) \frac{1}{165} + (-3) \frac{1}{165} = 0.$$

In general, if the pmf is symmetric around zero, the mean is zero.

## Problem 5

A school class of 120 students are driven in 3 buses to a symphonic performance. There are 36 students in bus 1, 40 in bus 2, and 44 in bus 3. One student is chosen at random. Let  $X$  denote the number of students on the bus of that randomly chosen student. What is  $E[X]$ ?

## Solution

We have  $P(X = 36) = 36/120$ ,  $P(X = 40) = 40/120$ , and  $P(X = 44) = 44/120$ . So

$$E[X] = (36) \frac{36}{120} + (40) \frac{40}{120} + (44) \frac{44}{120} = 40.2667$$

On the other hand, the average number of students on a bus is  $120/3 = 40$ . Why is the mean of the number of students in a bus of a randomly chosen student is larger than the average number of students on a bus?

## Problem 6

Let  $X$  be a random variable with possible values  $\{x_1, x_2, \dots\}$ . Let  $Y = g(X)$  and suppose the set of possible values for  $Y$  is  $\{y_1, y_2, \dots\}$ . Prove the Law Of The Unconscious Statistician,

$$E[Y] = \sum_i g(x_i) p_X(x_i).$$

## Solution

The LOTUS is nothing but a regrouping of probabilities:

$$\begin{aligned} E[Y] &= \sum_j y_j p_Y(y_j) \\ &= \sum_j y_j P(Y = y_j) \\ &= \sum_j y_j P(g(X) = y_j) \\ &= \sum_j y_j \sum_{i:g(x_i)=y_j} p_X(x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} y_j p_X(x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i) p_X(x_i) \end{aligned}$$

The summation  $\sum_j \sum_{i:g(x_i)=y_j}$  looks at every value in the set  $\{x_1, x_2, \dots\}$ , so

$$\begin{aligned} E[Y] &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i) p_X(x_i) \\ &= \sum_i g(x_i) p_X(x_i). \end{aligned}$$