STOCKHOLM UNIVERSITY
Department of Statistics

Michael Carlson
Emma Pettersson
2019-04-08

Solutions to Plenary Exercises: Plenary Exercise 1

Basic Statistics for Economists, 15 ECTS, STE101

EXERCISE 1

When rolling **one** dice, we can roll a 1, 2, 3, 4, 5 or 6. The sample space for a dice roll is then:

$$S = \{1,2,3,4,5,6\}$$

Assume that the dice is fair and that the probability of each outcome is equal (1/6).

Define the events E (even) and O (odd) as:

$$E = \text{“even number”} = \{2,4,6\} \quad ; \quad O = \text{“odd number”} = \{1,3,5\}$$

We can already illustrate the classical probability definition:

$$P(O) = \frac{size(O)}{size(S)} = \frac{3}{6} = \frac{1}{2} = 1 - P(O) = 1 - P(E) = P(E)$$

$$\boldsymbol{P(O) = P(E) = \frac{1}{2}}$$

It is reasonable to believe that the results from the three dice are independent of one another. The multiplication rule for independent events can give us the probability of a certain sequence of even and odd outcomes, and regardless of the exact sequence, the probability is always:

$$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

We can now define a new sample space for the experiment “throw three dice”. We can list all possible and equally likely outcomes, and group them according to how many even and odd numbers the outcomes have.

Possible Outcomes	P	No. Odd numbers	No. Even Numbers	
(EEE)	1/8	0	$3-0=3$	One occurance
(EEO) (EOE) (OEE)	1/8 1/8 1/8	1	$3-1=2$	Three occurances
(OOE) (OEO) (EOO)	1/8 1/8 1/8	2	$3-2=1$	Three occurances
(OOO)	1/8	3	$3-3=0$	One occurance

a. $A =$ " At least two odd"$= \{(OOE), (OEO), (EOO), (OOO)\}$, four out of eight outcomes.

$$P(A) = \frac{size(A)}{size(S)} = \frac{4}{8} = \mathbf{\frac{1}{2}}$$

b. $B =$ "At most two odd"= "Not three odd" $= S - \{(OOO)\}$, use the complement rule.

$$P(B) = 1 - P(OOO) = 1 - \frac{1}{8} = \mathbf{\frac{7}{8}}$$

c. $C =$ "Exactly one odd" $= \{(EEO),(EOE), (OEE)\}$, three out of eight outcomes.

$$P(C) = \mathbf{\frac{3}{8}}$$

d. $D =$ "Exactly two odd" $= \{(OOE), (OEO), (EOO)\}$, three out of eight outcomes

$$P(D) = \mathbf{\frac{3}{8}}$$

EXERCISE 2

Define the outcomes: $F_1 =$ first child is a girl, and $F_2 =$ second child is a girl.

Assume that F_1 and F_2 are independent, and that the events are equally likely:

$$P(F_1) = P(F_2) = 0.5$$

It follows that:

$$P(F_1 \cap F_2) = P(F_1) \times P(F_2)$$

a. The probability that both children are girls $= P(F_1 \cap F_2)$

$$P(F_1 \cap F_2) = P(F_1) \times P(F_2) = 0.5^2 = \mathbf{0.25}$$

Using the multiplication rule (the method used in (a)), we can set up a table with probabilities. We see that the probabilities of all the combined outcomes are the same:

	F_2	$\bar{F}_2$	
F_1	0.25	0.25	0.5
$\bar{F}_1$	0.25	0.25	0.5
	0.5	0.5	

b. We want to find the probability that both children are girls, <u>given</u> that we know at least one is a girl.

$$P(\text{both are girls} \mid \text{at least one girl}) = P(F_1 \cap F_2 | F_1 \cup F_2)$$

The probability that at least one child is a girl is $P(F_1 \cup F_2)$. This is the probability that the either the first child is a girl, the second child is a girl or that both are girls.

According to the addition rule:

$$P(F_1 \cup F_2) = P(F_1) + P(F_2) - P(F_1 \cap F_2) = 0.5 + 0.5 - 0.25 = 0.75$$

You can also use the complement rule. The complement of "at least one girl" is "both are not girls"

$$P(\text{at least one girl}) = 1 - P(\text{both are not girls}) = 1 - P(\bar{F}_1 \cap \bar{F}_2) = 1 - 0.25 = 0.75$$

Now we can find the conditional probability that both children are girls given that at least one is a girl.

$$P(\text{both are girls} \mid \text{at least one girl}) = P(F_1 \cap F_2 | F_1 \cup F_2) = \frac{P(F_1 \cap F_2)}{P(F_1 \cup F_2)} = \frac{0.25}{0.75} = \frac{\mathbf{1}}{\mathbf{3}}$$

A simpler way to look at the problem is to once again analyse the table. First, we take away the event $\overline{F_1} \cap \overline{F_2}$ from the sample space. If we know that at least one child is a girl, then we cannot include the event that both children are not girls. We now recalculate the probability of $F_1 \cap F_2$ in the new sample space.

	F_2	$\overline{F}_2$
F_1	0.25	0.25
$\overline{F}_1$	0.25	~~0.25~~

$$\frac{0.25}{0.25 + 0.25 + 0.25} = \frac{0.25}{0.75} = \frac{\mathbf{1}}{\mathbf{3}}$$

c. Recall that F_1 was defined as the event where the oldest child is a girl. The conditional probability that both children are girls given that the oldest child is a girl can be written as:

$$P(F_1 \cap F_2 | F_1)$$

In the table, we can take away any event that involves $\overline{F}_1$, since we are now only interested in events where F_1 occur. Then from the new sample space, we calculate the probability of $F_1 \cap F_2$.

	F_2	$\overline{F}_2$
F_1	0.25	0.25
$\overline{F}_1$	~~0.25~~	~~0.25~~

$$P(F_1 \cap F_2 | F_1) = \frac{0.25}{0.25 + 0.25} = \frac{0.25}{0.5} = \frac{\mathbf{1}}{\mathbf{2}}$$

Actually, since we assumed that the genders of the children are independent of each other, the probability that both are girls given that the first child is a girl, it's just the probability that the second child is a girl:

$$P(F_1 \cap F_2 | F_1) = \frac{P\big((F_1 \cap F_2) \cap F_1\big)}{P(F_1)} = \frac{P(F_1 \cap F_2)}{P(F_1)} = \frac{P(F_1) \cdot P(F_2)}{P(F_1)} = \frac{\cancel{P(F_1)} \cdot P(F_2)}{\cancel{P(F_1)}}$$

$$= P(F_2) = \frac{\mathbf{1}}{\mathbf{2}}$$

EXERCISE 3

Let M_1 = wins first game, and M_2 = wins second game.

$$P(M_1) = 0.1 \;\; ; \;\; P(M_2) = 0.3 \;\; ; \;\; P(\overline{M_1} \cap \overline{M_2}) = 0.65$$

The probability that the team wins exactly one match is the probability that they win either the first of second match, but not both. We can express this mathematically:

$$P(\text{win one game}) = P(M_1 \cup M_2) - P(M_1 \cap M_2)$$

Let's fill out a contingency table to help us solve the problem.

1. The first step is to fill in the information we got from the question (highlighted in grey).
2. The second step is to fill in the margins (highlighted in blue), remember that the sum of the row/column margins must equal 1.
3. The last step is to fill in the joint probabilities (highlighted in green).

	M_2	$\overline{M_2}$	
M_1	$0.3 - 0.25$ $= \mathbf{0.05}$	$0.7 - 0.65$ $= \mathbf{0.05}$	0.1
$\overline{M_1}$	$0.9 - 0.65$ $= \mathbf{0.25}$	0.65	$1 - 0.1 = \mathbf{0.9}$
	0.3	$1 - 0.3 = \mathbf{0.7}$	1.00

Now we can find the answer (Note that $P(M_1 \cup M_2) = P(M_1) + P(M_2) - P(M_1 \cap M_2)$):

$$P(M_1 \cup M_2) - P(M_1 \cap M_2) = P(M_1) + P(M_2) - 2P(M_1 \cap M_2) = 0.3 + 0.1 - 2(0.05) = \mathbf{0.3}$$

Another way to solve this problem is to analyse the contingency table and extract the probabilities where exactly one match is won:

$$P(Exactly\ one\ win) = P(M_1 \cap \overline{M_2}) + P(\overline{M_1} \cap M_2) = 0.05 + 0.25 = \mathbf{0.3}$$

Are the matches independent? No. $P(M_1 \cap M_2) \neq P(M_1)P(M_2)$

EXERCISE 4

Let C = gets cancer and let S = smoker.

The first thing we should do is extend the table. We do this by summing the joint probabilities by row/column to get the marginal probabilities.

	Gets cancer (C)	Does not get cancer ($\bar{C}$)	
Smoker (S)	0.0054	0.1546	0.1600
Non-Smoker($\bar{S}$)	0.0006	0.8394	0.8400
	0.0060	0.9940	1

a.
$$P(C|S) = \frac{P(C \cap S)}{P(S)} = \frac{0.0054}{0.1600} = 0.03375$$
Approximately 3.4% of smokers get cancer.

b.
$$P(C|\bar{S}) = \frac{P(C \cap \bar{S})}{P(\bar{S})} = \frac{0.0006}{0.8400} = 0.000714$$
Approximately 0.07% of non-smokers get cancer.

c. If you smoke, how many times larger are the chances of getting cancer?
To answer this question, we use our answers in a and b to calculate an odds ratio.
$$\frac{P(C|S)}{P(C|\bar{S})} = \frac{0.03375}{0.000714} \approx 47.27$$
The chances of getting cancer are about 47 times larger if you are a smoker.

EXERCISE 5

The expected value:

$$\mu_X = E(X) = \sum_{x \in S_x} xP(x) = 0 \cdot P(0) + 1 \cdot P(1) = 0(0.3) + 1(0.7) = \mathbf{0.7}$$

The variance:

$$\sigma_X^2 = Var(X) = \sum_{x \in S_x} (x - \mu_x)^2 P(x) = (0 - 0.7)^2 \cdot 0.3 + (1 - 0.7)^2 \cdot 0.7$$
$$= (0.49 \cdot 0.3) + (0.09 \cdot 0.7) = \mathbf{0.21}$$

Or alternatively:

$$\sigma_X^2 = Var(X) = E(X^2) - E(X)^2 = \sum_{x \in S_x} x^2 P(x) - \mu_X^2 = 0^2 \cdot 0.3 + 1^2 \cdot 0.7 - 0.7^2 = 0.7 - 0.7^2$$
$$= \mathbf{0.21}$$

You can also set up a table for your calculations:

x	0	1	Σ
$P(x)$	0.3	0.7	1.0
$xP(x)$	0	0.7	$0.7 = \mu_X$
$x - \mu_X$	-0.7	0.3	
$(x - \mu_X)^2$	0.49	0.09	
$(x - \mu_X)^2 P(x)$	0.147	0.063	$0.21 = \sigma_x^2$
x^2	0	1	
$x^2 P(x)$	0	0.7	$0.7 = \Sigma x^2 xP(x)$

EXERCISE 6

To verify that $P(x)$ is a probability distribution we have to check that:

(1) $P(x) \geq 0$ for all $x \in S_x$ and

(2) $\sum_{x \in S_x} P(x) = 1$

x	$P(x)$	$xP(x)$	$x^2 P(x)$
0	$\frac{3-\lvert 2-0\rvert}{9} = \frac{3-2}{9} = \mathbf{\frac{1}{9}}$	0	0
1	$\frac{3-\lvert 2-1\rvert}{9} = \frac{3-1}{9} = \mathbf{\frac{2}{9}}$	$2/9$	$2/9$
2	$\frac{3-\lvert 2-2\rvert}{9} = \frac{3-0}{9} = \mathbf{\frac{3}{9}}$	$6/9$	$12/9$
3	$\frac{3-\lvert 2-3\rvert}{9} = \frac{3-1}{9} = \mathbf{\frac{2}{9}}$	$6/9$	$18/9$
4	$\frac{3-\lvert 2-4\rvert}{9} = \frac{3-2}{9} = \mathbf{\frac{1}{9}}$	$4/9$	$16/9$
	$\sum_{x \in S_x} P(x) = \frac{9}{9} = 1$	$\sum_{x \in S_x} xP(x) = 2$	$\sum_{x \in S_x} x^2 P(x) = \frac{48}{9}$

We can confirm that the sum of the probabilities equals 1 and that $P(x) \geq x$ for all x.

$$E(X) = \sum_{x \in S_x} xP(x) = \frac{0+2+6+6+4}{9} = \mathbf{2}$$

$$Var(X) = \sum_{x \in S_x} x^2 P(x) - \mu_X^2 = \frac{0+2+12+18+16}{9} - 2^2 = \frac{48}{9} - 4 = \mathbf{\frac{12}{9}}$$

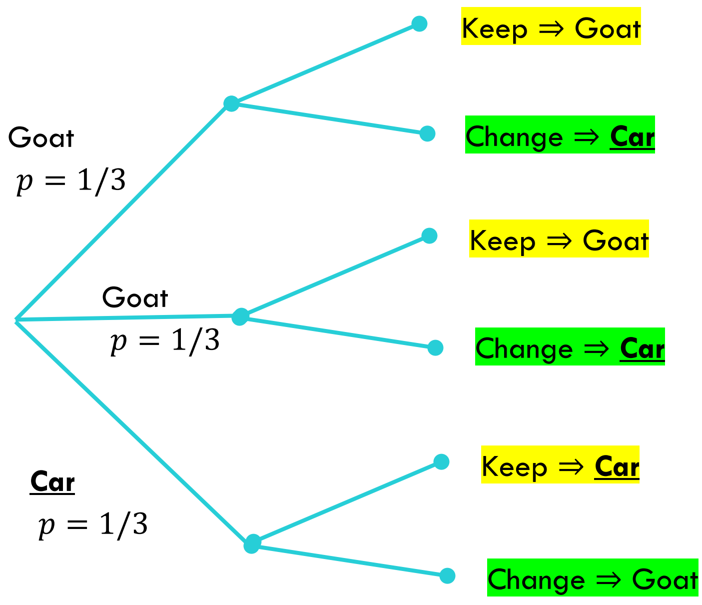
EXERCISE 7

Assume that the player chooses a door randomly, the probability of choosing any one door is then $1/3$.

The probability that the player selects the winning door, with the car behind it, is $1/3$. And that the probability of choosing a wrong door is $2/3$.

Remember that the program host always opens a door and reveals one of the goats.

Below we can illustrate the different outcomes, and the probabilities of the outcomes, in a tree diagram. We start at the left of the diagram, where a player chooses (with equal probability) one of three doors. Then the player is faced with a choice, should he change doors, or should he stick with his original decision? The strategy of changing doors is in yellow, and the strategy of keeping the same door in green.



We see that for a player who always changes doors, the probability of winning a car is $2/3$. While for a player who has the strategy to keep their original door, the probability of winning is $1/3$. In other words, changing doors is the better strategy.

Thought experiment: What if in the initial step you had 100 doors to choose from? Behind 99 doors there are goats and behind 1 door there is a car. You randomly choose one door (what is the probability that you chose the car?) and then the game show host opens 98 doors revealing a goat behind each. Would you change or keep?