Conditional Probability 1 – Solutions
COR1-GB.1305 – Statistics and Data Analysis

Compound Events and the Additive Rule

1. Suppose you roll two dice. Describe the event "you get a 1 on at least one of the dice."

Solution:

$$E = \{\text{ you get a 1 on the first die }\} \cup \{\text{ you get a 1 on the second die }\}.$$

2. Compute the probability of the event in problem 1.

Solution: Using the additive rule,

$$\begin{aligned} \mathrm{P}(E) &= \mathrm{P}(1 \text{ on the first}) + \mathrm{P}(1 \text{ on the second}) - \mathrm{P}(1 \text{ on the first } \textit{and } 1 \text{ on the second}) \\ &= \tfrac{1}{6} + \tfrac{1}{6} - \tfrac{1}{36} \\ &= \tfrac{11}{36}. \end{aligned}$$

3. Suppose that two customers give ratings to the same restaurant on Yelp.
   (a) Express the event "the average of their ratings is 3.5 or 4" as a union of two other events. *Hint: this is the same event as "the sum of their ratings is 7 or 8."*

Solution: Define two events:

$$\begin{aligned} S_7 &= \{\text{ the sum of their ratings is 7 }\} \\ &= \{(2,5),(3,4),(4,3),(5,2)\}, \\ S_8 &= \{\text{ the sum of their ratings is 8 }\} \\ &= \{(3,5),(4,4),(5,3)\}. \end{aligned}$$

Then, the event we care about is $A = S_7 \cup S_8$.

   (b) Suppose that both customers randomly assign their ratings with equal probability for all possible star ratings, so that all 25 sample points have equal probability. Compute the probability of the event in part (a).

Solution: Using the additive rule,

$$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}(S_7 \cup S_8) \\ &= \mathrm{P}(S_7) + \mathrm{P}(S_8) - \mathrm{P}(S_7 \cap S_8). \end{aligned}$$

We note that the sum can't be 7 and 8 simultaneously, so S_7 and S_8 are mutually exclusive events, i.e. $S_7 \cap S_8 = \emptyset$. Thus,

$$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}(S_7) + \mathrm{P}(S_8) \\ &= \tfrac{4}{25} + \tfrac{3}{25} \\ &= \tfrac{7}{25} \\ &= 28\%. \end{aligned}$$

Complementary Events and the Complement Rule

4. Suppose that 60% of NYU MBA students own iPhones. If you pick a random NYU MBA student, what is the probability that he or she will *not* own an iPhone?

Solution: The sample space is the set of all students. Let

$$A = \{ \text{ the randomly picked student owns an iPhone } \}.$$

Then,

$$A^c = \{ \text{ the randomly picked student does not own an iPhone } \},$$

so by the complement rule,

$$\begin{aligned} \mathrm{P}(A^c) &= 1 - \mathrm{P}(A) \\ &= 1 - .60 \\ &= .40. \end{aligned}$$

5. Suppose you flip five coins. What is the probability of getting at least one head?
Hint: what is the complement of this event?

Solution: The sample space, Ω, is the set of all possible outcomes for the five flips. Since there are 5 independent flips, and each has 2 possible outcomes, we have that $|\Omega| = 2^5 = 32$.

Let

$$A = \{ \text{ you get at least one head } \}.$$

Then,

$$\begin{aligned} A^c &= \{ \text{ you don't get any heads } \} \\ &= \{(T, T, T, T, T)\}. \end{aligned}$$

Thus, by the complement rule,

$$\begin{aligned} \mathrm{P}(A) &= 1 - \mathrm{P}(A^c) \\ &= 1 - \tfrac{1}{32} \\ &= \tfrac{31}{32}. \end{aligned}$$

Conditional Probability

6. Here is a table of the tabulated frequencies for the expected starting salary and gender for the respondents to the class survey.

	Gender		
Salary ($1K)	Female	Male	Total
$(0, 100]$	11	4	15
$(100, 125]$	5	7	12
$(125, \infty]$	5	15	20
Total	21	26	47

(a) Express the following statements as conditional probabilities:
- $\frac{11}{21} \approx 52\%$ of the females listed a starting salary of \$100K or lower.
- $\frac{11}{15} \approx 73\%$ of those listing starting salaries of \$100K or lower are female.

Solution:

$$P(\text{Salary} \leq \$100\text{K} \mid \text{Female}) = \frac{11}{21},$$
$$P(\text{Female} \mid \text{Salary} \leq \$100\text{K}) = \frac{11}{15}.$$

(b) Compute $P(\text{Male} \mid \text{Salary} > \$125\text{K})$ and $P(\text{Salary} > \$125\text{K} \mid \text{Male})$. Explain the difference between these two quantities.

Solution:

$$P(\text{Male} \mid \text{Salary} > \$125\text{K}) = \frac{15}{20} = 75\%,$$
$$P(\text{Salary} > \$125\text{K} \mid \text{Male}) = \frac{15}{26} \approx 58\%.$$

The quantity $P(\text{Male} \mid \text{Salary} > \$125\text{K})$ is the proportion of those listing salaries above \$125K that are male. The quantity $P(\text{Salary} > \$125\text{K} \mid \text{Male})$ is the proportion of males listing salary above \$125K.

7. The following table lists the pick-up and drop-off locations of approximately 170 million yellow cab taxi trips made in New York City in 2013. Numbers are reported in thousands.

	Drop-off					
Pick-up	Bronx	Brooklyn	Manhattan	Queens	Staten Is.	Total
Bronx	53	1	37	4	0	95
Brooklyn	8	2,707	1,598	273	2	4,588
Manhattan	638	5,458	143,656	5,906	22	155,680
Queens	122	1,022	5,058	2,281	8	8,491
Staten Is.	0	0	0	0	3	3
Total	821	9,188	150,349	8,464	35	168,857

(a) Find $P(\text{drop-off Brooklyn} \mid \text{pick-up Manhattan})$ and $P(\text{pick-up Manhattan} \mid \text{drop-off Brooklyn})$. Explain the difference between these two quantities.

Solution:

$$P(\text{drop-off Brooklyn} \mid \text{pick-up Manhattan}) = \frac{5458}{155680} \approx 3.5\%,$$
$$P(\text{pick-up Manhattan} \mid \text{drop-off Brooklyn}) = \frac{5458}{9188} \approx 59.4\%.$$

3.5% of the rides that pick up in Manhattan drop off in Brooklyn; 59.4% of the rides that drop off in Brooklyn originate in Manhattan.

(b) Express the following statement as a conditional probability: "29% of the trips with drop-off locations in Brooklyn originated in the same borough."

Solution:

$$P(\text{pick-up Brooklyn} \mid \text{drop-off Brooklyn}) = \frac{2707}{9188} = 29\%.$$

Note:

$$P(\text{drop-off Brooklyn} \mid \text{pick-up Brooklyn}) = \frac{2707}{4588} = 59\%.$$

The Multiplicative Rule

8. Out of the 58 students enrolled in the class, 24 are female (41%) and 34 are male (59%). Suppose that we randomly select two different students.

   (a) What is the probability that both students are male?

   Solution: Define the two events

   $$\begin{aligned} A &= \text{the first student picked is male} \\ B &= \text{the second student picked is male.} \end{aligned}$$

   Then, $P(A) = \frac{34}{58}$, and $P(B \mid A) = \frac{33}{57}$. Thus, the probability that both will be male is

   $$\begin{aligned} P(A \cap B) &= P(A)P(B \mid A) \\ &= \frac{34}{58} \cdot \frac{33}{57} \\ &= \frac{1122}{3306} \\ &\approx 34\%. \end{aligned}$$

   (b) What is the probability that both students are female?

   Solution: Using the events A and B defined in the previous part, $P(A^c) = \frac{24}{58}$ and $P(B^c \mid A^c) = \frac{23}{57}$. Thus, the probability that both will be female is

   $$\begin{aligned} P(A^c \cap B^c) &= P(A^c)P(B^c \mid A^c) \\ &= \frac{24}{58} \cdot \frac{23}{57} \\ &= \frac{552}{3306} \\ &\approx 17\%. \end{aligned}$$

   (c) What is the probability that one of the students is male and one of the students is female?

   Solution: The event "one student is male and the other is female" is equivalent to the compound event $(A \cap B^c) \cup (A^c \cap B)$; that is, either the first is male and the second is female, or the first is female and the second is male. Since $A \cap B^c$ and $A^c \cap B$ are mutually exclusive, it follows that

   $$P(\text{one male and one female}) = P(A \cap B^c) + P(A^c \cap B).$$

Using the multiplicative rule,

$$\begin{aligned} P(A \cap B^c) &= P(A)P(B^c \mid A) \\ &= \frac{34}{58}\frac{24}{57} \\ &= \frac{816}{3306} \end{aligned}$$

$$\begin{aligned} P(A^c \cap B) &= P(A^c)P(B \mid A^c) \\ &= \frac{24}{58}\frac{34}{57} \\ &= \frac{816}{3306}. \end{aligned}$$

Thus,

$$\begin{aligned} P(\text{one male and one female}) &= \frac{816}{3306} + \frac{816}{3306} \\ &= \frac{1632}{3306} \\ &\approx 49\%. \end{aligned}$$

9. Of the 48 students who filled out the survey, 33 indicated that they drink at least one cup of coffee per day, while 15 indicated that they do not drink coffee on a typical day. Suppose that we randomly select two different survey respondents.

   (a) What is the probability that both students regularly drink coffee?

   Solution:

   $$\frac{33}{48} \cdot \frac{32}{47} = \frac{1056}{2256} \approx 47\%.$$

   (b) What is the probability that neither student regularly drinks coffee?

   Solution:

   $$\frac{15}{48} \cdot \frac{14}{47} = \frac{210}{2256} \approx 9\%.$$

   (c) What is the probability that exactly one student regularly drinks coffee?

   Solution:

   $$\frac{33}{48} \cdot \frac{15}{47} + \frac{15}{48} \cdot \frac{33}{47} = \frac{990}{2256} \approx 44\%.$$