Probability – Solutions
STAT-UB.0001 – Statistics for Business Control

Sample Points and Sample Spaces

1. In the following two experiments, what are the sample points and the sample space?

   (a) You flip a coin.

   Solution: The sample points are H, "the outcome is heads," and T, "the outcome is tails." The sample space is the set of all sample points: $\Omega = \{H, T\}$.

   (b) You roll a 6-sided die.

   Solution: The sample points are the possible outcomes of the die: 1, 2, 3, 4, 5, 6. The sample space is the set of all sample points: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

2. Suppose that a customer visits a restaurant and leaves a review on Yelp with 1–5 stars. What are the sample points and the sample space for the customer's star rating?

   Solution: The sample points are the possible star ratings: 1, 2, 3, 4, 5. The sample space is the set of all sample points: $\Omega = \{1, 2, 3, 4, 5\}$.

3. Suppose that two customers visit a restaurant, and that they both leave Yelp reviews with 1–5 stars each. What are the sample points and the sample space for the pair of star ratings?

   Solution: Each sample point can be represented by an ordered pair (i, j), where i is the first customer's star rating, and j is second customer's star rating. The sample points are the elements of the following table.

	1	2	$\cdots$	5
1	(1,1)	(1,2)	$\cdots$	(1,5)
2	(2,1)	(2,2)	$\cdots$	(2,5)
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
5	(5,1)	(5,2)	$\cdots$	(5,5)

   The sample space is the set of all 25 sample points: $\Omega = \{(1, 1), (1, 2), \ldots, (5, 5)\}$.

4. Suppose you randomly pick a respondent from the class survey, then record their undergraduate major and gender. What are the sample points and the sample space? Assume that major is either "Business," "Humanities/Social Science," or "Science/Engineering."

> **Solution:** Each sample point is a gender-major pair. The sample space is the set of all possible gender-major pairs: $\Omega = \{(\text{Bus.}, \text{Female})$, $(\text{Bus.}, \text{Male})$, $(\text{Hum./Soc. Sci.}, \text{Female})$, $(\text{Hum./Soc. Sci.}, \text{Male})$, $(\text{Sci./Eng.}, \text{Female})$, $(\text{Sci./Eng.}, \text{Male})\}$. Note that the sample space is all *possible* gender-major pairs, not all *observed* gender-major pairs.

Events

5. Suppose that a customer leaves a Yelp rating (1–5 stars) for a restaurant. Describe the event "the rating is odd (not even)."

> **Solution:**
> $$O = \{1, 3, 5\}.$$

6. Suppose you randomly pick a respondent from the class survey, then record their undergraduate major and gender. Assume that undergraduate major is listed as "Business", "Hum./Soc. Sci.", or "Sci./Eng.", and that gender is listed as "Male" or "Female".

   (a) List the sample points in the event "the major is Business."

   > **Solution:** $\text{Business} = \{(\text{Business}, \text{Female}), (\text{Business}, \text{Male})\}$.

   (b) List the sample points in the event "the gender is Male."

   > **Solution:** $\text{Male} = \big\{\big(\text{Business}, \text{Male}\big), \big(\text{Hum./Soc. Sci.}, \text{Male}\big), \big(\text{Sci./Eng.}, \text{Male}\big)\big\}$.

Probability

7. Suppose you randomly pick a respondent from the class survey and record their undergraduate major and gender.

   (a) Use the following table of recorded survey response frequencies to compute the probabilities of the sample points.

	Gender		
Undergrad Major	Female	Male	Total
Business	9	6	15
Hum./Soc. Sci.	10	12	22
Sci./Eng.	2	8	10
Total	21	26	47

Solution: To compute the probabilities for the 6 sample points corresponding to the cells of the table, we take the recorded frequency and divide by the total number of survey respondents. We have

$$\begin{aligned} \mathrm{P}\left(\text{Bus., Female}\right) &= \frac{9}{47} \approx .19, \\ \mathrm{P}\left(\text{Bus., Male}\right) &= \frac{6}{47} \approx .13, \\ \mathrm{P}\left(\text{Hum./Soc. Sci., Female}\right) &= \frac{10}{47} \approx .21, \\ \mathrm{P}\left(\text{Hum./Soc. Sci., Male}\right) &= \frac{12}{47} \approx .26, \\ \mathrm{P}\left(\text{Sci./Eng., Female}\right) &= \frac{2}{47} \approx .04, \\ \mathrm{P}\left(\text{Sci./Eng., Male}\right) &= \frac{8}{47} \approx .17. \end{aligned}$$

   (b) Find the probability that the undergraduate major is Business.

Solution:

$$\begin{aligned} \mathrm{P}(\text{Business}) &= \tfrac{9}{47} + \tfrac{6}{47} \\ &= \tfrac{15}{47} \\ &\approx 32\%. \end{aligned}$$

(c) Find the probability that the gender is Male.

Solution:

$$\begin{aligned} \mathrm{P}(\text{Male}) &= \tfrac{6}{47} + \tfrac{12}{47} + \tfrac{8}{47} \\ &= \tfrac{26}{47} \\ &\approx 55\%. \end{aligned}$$

(d) Find the probability the undergraduate major is Humanities/Social Science.

Solution:

$$\begin{aligned} \mathrm{P}(\text{Hum./Soc. Sci.}) &= \tfrac{10}{47} + \tfrac{12}{47} \\ &= \tfrac{22}{47} \\ &\approx 47\%. \end{aligned}$$

8. Suppose that a customer's Yelp rating is random, and that the probabilities for the possible star ratings are $p_1 = 10\%$, $p_2 = 30\%$, $p_3 = 25\%$, $p_4 = 20\%$, $p_5 = 15\%$. Find the probability that the rating is odd.

Solution: We add up the probabilities of the sample points in the event:

$$\begin{aligned} \mathrm{P}(\{1,3,5\}) &= p_1 + p_3 + p_5 \\ &= 10\% + 25\% + 15\% \\ &= 50\%. \end{aligned}$$

Compound Events and the Additive Rule

9. Suppose you pick a random survey respondent and record their undergraduate major and gender.

   (a) List the sample points in the event "the major is Business or the gender is Male."

   Solution: Denote the event by A. Then,

   $$A = \{(\text{Business}, \text{Female}), (\text{Business}, \text{Male}), (\text{Hum./Soc. Sci.}, \text{Male}), (\text{Sci./Eng.}, \text{Male})\}.$$

   (b) Compute the probability of the event in part (a) by adding the probabilities of all of the sample points in the event.

   Solution:

   $$\begin{aligned} \mathrm{P}(A) &= \tfrac{9}{47} + \tfrac{6}{47} + \tfrac{12}{47} + \tfrac{8}{47} \\ &= \tfrac{35}{47} \\ &\approx 74\%. \end{aligned}$$

   (c) Express the event "the major is Business or the gender is Male" as a union of two other events.

   Solution:

   $$\begin{aligned} A &= \{\text{major is Business or gender is Male}\} \\ &= \text{Business} \cup \text{Male}. \end{aligned}$$

   (d) Compute the probability of the event using the additive rule.

   Solution:

   $$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}(\text{Business} \cup \text{Male}) \\ &= \mathrm{P}(\text{Business}) + \mathrm{P}(\text{Male}) - \mathrm{P}(\text{Business} \cap \text{Male}) \\ &= \tfrac{15}{47} + \tfrac{26}{47} - \tfrac{6}{47} \\ &= \tfrac{35}{47} \\ &\approx 74\%. \end{aligned}$$

10. Suppose that two customers give ratings (1–5 stars) to the same restaurant on Yelp.

   (a) Express the event "at least one customer gives a 1 star rating" as a union of two other events.

   > **Solution:**
   >
   > $$A = \{ \text{ the first customer gives a 1 star rating } \} \\ \cup \{ \text{ the second customer gives a 1 star rating } \}.$$

   (b) Suppose that both customers randomly assign their ratings, giving equal probabilities to all possible star ratings. In this case, all 25 sample points have equal probability. Compute the probability of the event in part (a).

   > **Solution:** Using the additive rule,
   >
   > $$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}(1 \text{ from first customer}) + \mathrm{P}(1 \text{ from second customer}) \\ &\quad - \mathrm{P}(1 \text{ from the first customer } \textit{and} \text{ 1 from second customer}) \\ &= \tfrac{1}{5} + \tfrac{1}{5} - \tfrac{1}{25} \\ &= \tfrac{9}{25} \\ &= 36\%. \end{aligned}$$

11. 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) As in problem 10(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

12. Suppose that 60% of NYU undergraduates own iPhones. If you pick a random NYU undegraduate, 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,

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

13. 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

14. Here is a table of the tabulated frequencies for undergrad major and gender for the respondents to a class survey.

	Gender		
Undergrad Major	Female	Male	Total
Business	9	6	15
Hum./Soc. Sci.	10	12	22
Sci./Eng.	2	8	10
Total	21	26	47

(a) Express the following statements as conditional probabilities:

- $\frac{9}{21} \approx 43\%$ of the females have undergrad major in Business.
- $\frac{9}{15} = 60\%$ of those having undergrad major in Business are female.

Solution:

$$P(\text{Business} \mid \text{Female}) = \frac{9}{21},$$
$$P(\text{Female} \mid \text{Business}) = \frac{9}{15}.$$

(b) Compute $P(\text{Male} \mid \text{Sci./Eng.})$ and $P(\text{Sci./Eng.} \mid \text{Male})$. Explain the difference between these two quantities.

Solution:

$$P(\text{Male} \mid \text{Sci./Eng.}) = \frac{8}{10} = 80\%,$$
$$P(\text{Sci./Eng.} \mid \text{Male}) = \frac{8}{26} \approx 31\%.$$

The quantity $P(\text{Male} \mid \text{Sci./Eng.})$ is the proportion of those having undergrad major in Sci./Eng. that are male. The quantity $P(\text{Sci./Eng.} \mid \text{Male})$ is the proportion of males having undergrad major in Sci./Eng.

15. 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(drop-off Brooklyn | pick-up Manhattan) and P(pick-up Manhattan | 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

16. 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

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

    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,

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

17. 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:

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

18. A class has 20 students. What is the probability that at least two students have the same birthday? Assume that each person in the class was assigned a random birthday between January 1 and December 31.

Solution: Assume that everyone in the class is randomly assigned a birthday, which corresponds to number between 1 and 365 representing the day of the year. It turns out to be much easier to compute the probability using the complement rule, as

$$\text{P(at least 2 people have the same birthday)} = 1 - \text{P(all 20 birthdays are different)}.$$

The next trick is to write the event that all 50 birthdays are different in a redundant way:

$$\begin{aligned}\{\text{all 50 birthdays are different}\} &= \{\text{first 2 are different}\} \cap \{\text{first 3 are different}\} \\ &\cap \{\text{first 4 are different}\} \cap \{\text{first 5 are different}\} \cap \cdots \cap \{\text{first 50 are different}\}.\end{aligned}$$

In class we showed how to use the multiplicative rule repeatedly to get:

$$\begin{aligned}\text{P(all 50 birthdays are different)} &= \text{P(first 2 diff.)} \cdot \text{P(first 3 diff.} \mid \text{first 2 diff.)} \\ &\cdot \text{P(first 4 diff.} \mid \text{first 3 diff.)} \cdot \text{P(first 5 diff.} \mid \text{first 4 diff.)} \cdot \cdots \cdot \text{P(first 50 diff.} \mid \text{first 49 diff.)}.\end{aligned}$$

Using this expression we can compute

$$\text{P(all 50 birthdays are different)} = \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \cdots \cdot \frac{365-49}{365}.$$

We can do a similar calculation for other class sizes. The following table shows the probabilities of having at least two students with the same birthday for various class sizes:

Class Size	P(all diff.)	P(at least 2 same)
10	88%	12%
20	59%	41%
30	29%	71%
40	11%	89%
50	3%	97%
60	0.6%	99.4%
70	0.08%	99.92%