UNIVERSITY OF TEXAS AT AUSTIN, DEPARTMENT OF MATHEMATICS
M358K - Applied Statistics

THE REAL IN-TERM ONE PROBLEM BANK

1.1. The binomial distribution in R.

Problem 1.1. (5 points) What is the **R** output of the following command:

```
>dbinom(4,5,0.2)
```

(a) 0.0064
(b) 0.064
(c) 0.64
(d) 0.9875
(e) None of the above.

Solution: (a)
This is exactly the probability that a random variable $X \sim Binomial(size = 5, p = 0.2)$ takes the value equal to 4. We have

$$\mathbb{P}[X = 4] = \binom{5}{4}\left(\frac{1}{5}\right)^4\left(\frac{4}{5}\right)^1 = \frac{4}{5^4} = 0.0064.$$

Problem 1.2. (5 points) What is the **R** output of the following command:

```
>dbinom(1,4,0.2)
```

(a) 0.0064
(b) 0.064
(c) 0.4096
(d) 0.9875
(e) None of the above.

Solution: (c)
This is exactly the probability that a random variable $X \sim Binomial(size = 4, p = 0.2)$ takes the value equal to 1. We have

$$\mathbb{P}[X = 1] = \binom{4}{1}\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^3 = \frac{4^4}{5^4} = 0.4096.$$

Problem 1.3. (5 points) What is the **R** output of the following command:

```
>pbinom(4,5,0.5)
```

(a) 0.375
(b) 0.5
(c) 0.96875
(d) 0.9875
(e) None of the above.

Solution: (c)
This is exactly the probability that a random variable $X \sim Binomial(size = 5, p = 0.5)$ takes the value less than or equal to 4. We have

$$\mathbb{P}[X \leq 4] = 1 - \mathbb{P}[X = 5] = 1 - \binom{5}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 = 1 - \left(\frac{1}{2}\right)^5 = 0.96875.$$

Problem 1.4. (5 points) What is the **R** output of the following command:

```
>pbinom(2,3,0.5)
```

(a) 0.375
(b) 0.5
(c) 0.6875
(d) 0.875
(e) None of the above.

Solution: (d)
This is exactly the probability that a random variable $X \sim Binomial(size = 3, p = 0.5)$ takes the value less than or equal to 2. We have

$$\mathbb{P}[X \leq 2] = 1 - \mathbb{P}[X = 3] = 1 - \binom{3}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 = 1 - \left(\frac{1}{2}\right)^3 = 0.875.$$

Problem 1.5. (5 points) What is the **R** output of the following command:

```
>dbinom(2,4,0.5)
```

(a) 0.375
(b) 0.5
(c) 0.6875
(d) 0.875
(e) None of the above.

Solution: (a)
This is exactly the probability that a random variable $X \sim Binomial(size = 4, p = 0.5)$ takes the value equal to 2. We have

$$\mathbb{P}[X = 2] = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{6}{16} = 0.375.$$

1.2. **Independent events.**

Problem 1.6. (5 points) Three independent fair coins are tossed. Let

$$E_1 = \{\text{the outcome on coin \#1 matches the outcome on coin \#2}\}$$
$$E_2 = \{\text{the outcome on coin \#2 matches the outcome on coin \#3}\}$$
$$E_3 = \{\text{the outcome on coin \#3 matches the outcome on coin \#1}\}$$

Which of the following statements is **not** correct?

(a) E_1, E_2 and E_3 are independent.
(b) E_1 and E_2 are independent.
(c) E_2 and E_3 are independent.
(d) E_3 and E_1 are independent.
(e) E_1 and E_3 are independent.

Solution: (a)

Problem 1.7. (5 points) In a class there are four first-year Thunderbirds, six first-year Wampuses, and six second-year Thunderbirds. How many second-year Wampuses must be present if house (Thunderbirds and Wampuses) and year (first and second) are to be independent when a student is selected at random? There are no other students in this class!

(a) 3
(b) 5
(c) 7
(d) 9
(e) None of the above.

Solution: (d)
Let the total number of students be n and let the number of second-year Wampuses be k. Then, $n = 16 + k$. One condition for independence reads as (in obvious notation)

$$\mathbb{P}[T \cap I] = \mathbb{P}[T] \times \mathbb{P}[I] \quad \Leftrightarrow \quad \frac{4}{n} = \frac{10}{n} \times \frac{10}{n} \quad \Leftrightarrow \quad n = 25.$$

We conclude that $k = 25 - 16 = 9$. One can easily check that all the other independence conditions hold as well:

$$\mathbb{P}[T \cap II] = \mathbb{P}[T] \times \mathbb{P}[II] \quad \Leftrightarrow \quad \frac{6}{25} = \frac{10}{25} \times \frac{15}{25} \quad \Leftrightarrow \quad TRUE,$$
$$\mathbb{P}[W \cap I] = \mathbb{P}[W] \times \mathbb{P}[I] \quad \Leftrightarrow \quad \frac{6}{25} = \frac{15}{25} \times \frac{10}{25} \quad \Leftrightarrow \quad TRUE,$$
$$\mathbb{P}[W \cap II] = \mathbb{P}[W] \times \mathbb{P}[II] \quad \Leftrightarrow \quad \frac{9}{25} = \frac{15}{25} \times \frac{15}{25} \quad \Leftrightarrow \quad TRUE.$$

Problem 1.8. (5 points) A coin is tossed, and, independently, a 6-sided die is rolled. Let

$$A = \{4 \text{ is obtained on the die}\} \text{ and}$$
$$B = \{\textit{Heads} \text{ is obtained on the coin and an even number is obtained on the die}\}.$$

Then

(a) A and B are mutually exclusive
(b) A and B are independent
(c) $A \subseteq B$
(d) $A \cap B = B$
(e) none of the above

Solution: The correct answer is (e).

Problem 1.9. (5 points) You generate an physically impossible five-sided fair die in your computer. Its outcomes are 1, 2, 3, 4, and 5. They are all equally likely. You "roll" your virtual die twice independently. Define these events:

$$A = \{\text{the total of two rolls is } 10\},$$
$$B = \{\text{the total of two rolls is } 7\},$$
$$C = \{\text{the second roll resulted in a strictly higher number}\}.$$

Which one of these statements is **correct**?

(a) A and B are independent.
(b) A and C are independent.
(c) B and C are independent.
(d) A and B and C are independent.
(e) None of the above.

Solution: (e)
We have the following probabilities:

$$\mathbb{P}[A] = \frac{1}{25}, \quad \mathbb{P}[B] = \frac{2}{25}, \quad \mathbb{P}[C] = \frac{10}{25} = \frac{2}{5},$$
$$\mathbb{P}[A \cap B] = \mathbb{P}[\emptyset] = 0, \quad \mathbb{P}[A \cap C] = \mathbb{P}[\emptyset] = 0,$$

Obviously, none of the combinations of these events are independent.

Problem 1.10. (5 points) *Source: Problem 3.1.9 from Pitman's "Probability".*
A box contains 8 balls. Two are red, two are yellow, two are green and two are purple.
Balls are drawn from the box **without replacement** until the color appears that has appeared before. Let X be the random variable denoting the number of draws that are made. Let $p_X(2) = \mathbb{P}[X = 2]$. Then,

(a) $p_X(2) = 7/35 = 1/5$
(b) $p_X(2) = 5/35 = 1/7$
(c) $p_X(2) = 4/35$
(d) $p_X(2) = 1/35$
(e) None of the above

Solution: (b)

$$\begin{aligned} p_X(2) &= \mathbb{P}[\{\text{first two balls are red}\}] + \mathbb{P}[\{\text{first two balls are yellow}\}] \\ &\quad + \mathbb{P}[\{\text{first two balls are green}\}] + \mathbb{P}[\{\text{first two balls are purple}\}] \\ &= 4 \cdot \frac{2}{8} \cdot \frac{1}{7} = \frac{1}{7}\,. \end{aligned}$$

Problem 1.11. (5 points) Let $\Omega = \{a_1, a_2, a_3, a_4\}$ be a sample space, and let $\mathbb{P}$ be a probability on Ω. Assume that $\mathbb{P}[\{a_2, a_3\}] = 2/3, \mathbb{P}[\{a_2, a_4\}] = 1/2$ and $\mathbb{P}[\{a_2\}] = 1/3$. Then we have that $\mathbb{P}[\{a_1\}]$ equals the following value:

(a) 1/12
(b) 1/6
(c) 1/3
(d) 1/2
(e) None of the above

Solution: (b)
From the given values of $\mathbb{P}$ on certain events, we conclude that

$$\mathbb{P}[\{a_3\}] = \mathbb{P}[\{a_2, a_3\}] - \mathbb{P}[\{a_2\}] = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}\,,$$
$$\mathbb{P}[\{a_4\}] = \mathbb{P}[\{a_2, a_4\}] - \mathbb{P}[\{a_2\}] = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}\,.$$

So,

$$\mathbb{P}[\{a_1\}] = 1 - (\mathbb{P}[\{a_2\}] + \mathbb{P}[\{a_3\}] + \mathbb{P}[\{a_4\}]) = \frac{1}{6}\,.$$

Problem 1.12. (5 points) Two dice are rolled. What is the probability that the maximum of the two outcomes equals 4?

(a) 1/4
(b) 3/10
(c) 13/28
(d) 7/36
(e) None of the above

Solution: (d)
Let X denote the outcome on the first die and Y the outcome on the second die. The probability we are looking for is

$$\begin{aligned} \mathbb{P}[\max(X,Y) = 4] &= \mathbb{P}[X = 4, Y = 4] + \mathbb{P}[X = 4, Y \in \{1,2,3\}] + \mathbb{P}[X \in \{1,2,3\}, Y = 4] \\ &= \frac{1}{36} + \frac{1}{6} \cdot \frac{1}{2} + \frac{1}{6} \cdot \frac{1}{2} \\ &= \frac{7}{36}\,. \end{aligned}$$

Problem 1.13. (5 points) Let $\Omega = \{a_1, a_2, a_3, a_4\}$ be an outcome space, and let $\mathbb{P}$ be a probability distribution on Ω. Assume that $\mathbb{P}[\{a_1\}] = 1/3, \mathbb{P}[\{a_2\}] = 1/6$ and $\mathbb{P}[\{a_3\}] = 1/9$. Then we have that $\mathbb{P}[\{a_4\}]$ equals the following value:

(a) 1/3
(b) 2/3
(c) 7/18
(d) 7/9
(e) None of the above

Solution: (c)
For any outcome space Ω, from the axioms of probability, we must have that $\mathbb{P}[\Omega] = 1$. In this case, $\Omega = \{a_1, a_2, a_3, a_4\}$, and so

$$\begin{aligned}\mathbb{P}[\Omega] &= \mathbb{P}[\{a_1, a_2, a_3, a_4\}] \\ &= \mathbb{P}[\{a_1\}] + \mathbb{P}[\{a_2\}] + \mathbb{P}[\{a_3\}] + \mathbb{P}[\{a_4\}] \\ &= \frac{1}{3} + \frac{1}{6} + \frac{1}{9} + \mathbb{P}[\{a_4\}].\end{aligned}$$

Hence,

$$\mathbb{P}[\{a_4\}] = 1 - \frac{11}{18} = \frac{7}{18}\,.$$

Problem 1.14. (5 pts) Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ be an outcome space, and let $\mathbb{P}$ be a probability distribution on Ω. Assume that $\mathbb{P}[\{\omega_1, \omega_2\}] = 1/3, \mathbb{P}[\{\omega_2, \omega_3\}] = 1/4$ and $\mathbb{P}[\{\omega_1, \omega_3\}] = 1/9$. Then we have that $\mathbb{P}[\{\omega_4\}]$ equals the following value:

(a) $1/4$
(b) $11/18$
(c) $7/36$
(d) $47/72$
(e) None of the above

Solution: (d)
For any outcome space Ω, from the axioms of probability, we must have that $\mathbb{P}[\Omega] = 1$. In this case, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, and so

$$\mathbb{P}[\Omega] = \mathbb{P}[\{\omega_1, \omega_2, \omega_3, \omega_4\}] = \mathbb{P}[\{\omega_1, \omega_2, \omega_3\}] + \mathbb{P}[\{\omega_4\}] = \frac{1}{2}\left(\frac{1}{3} + \frac{1}{4} + \frac{1}{9}\right) + \mathbb{P}[\{\omega_4\}].$$

Hence,

$$\mathbb{P}[\{\omega_4\}] = 1 - \frac{25}{72} = \frac{47}{72}\,.$$

Problem 1.15. (5 pts) *Source: Sample P exam problem set.*

An insurance company pays hospital claims.

The number of claims that include emergency room or operating room charges is 85% of the total number of hospital claims.

The number of claims that do not include emergency room charges is 25% of the total number of claims.

The occurrences of emergency room charges and operating room charges are independent.

Calculate the probability that a claim submitted to the insurance company includes operating room charges.

(a) 0.1
(b) 0.2
(c) 0.25
(d) 0.4
(e) None of the above

Solution: (d)

Let us denote by E the event that a given claim includes emergency room charges and by O the event that it includes operating room charges. Then, we can formalize the data provided in the problem as

$$\begin{aligned}\mathbb{P}[E \cup O] &= 0.85\\ \mathbb{P}[E^c] &= 0.25 \Rightarrow \mathbb{P}[E] = 0.75\\ \mathbb{P}[E \cap O] &= \mathbb{P}[E]\mathbb{P}[O].\end{aligned}$$

We need to find $\mathbb{P}[O]$.

From the first equality above, we get

$$0.85 = \mathbb{P}[E \cup O] = \mathbb{P}[E] + \mathbb{P}[O] - \mathbb{P}[E \cap O].$$

Then, using the second qnd third equalities,

$$0.85 = \mathbb{P}[E \cup O] = 0.75 + \mathbb{P}[O] - 0.75\mathbb{P}[O].$$

So, $\mathbb{P}[O] = 0.1/0.25 = 0.4$.

Problem 1.16. (5 points) In Country X, the latest census has revealed the following:

- 40% of the population exercise regularly,
- 30% own a dog,
- 20% like cauliflower,
- 60% of all dog owners exercise regularly,
- 10% own a dog and like cauliflower,
- 4% exercise regularly, own a dog and like cauliflower.

If it is known that the randomly selected person either likes cauliflower or owns a dog (or both), what is the probability that he/she exercises regularly, owns a dog and likes cauliflower?

(a) 0.08

(b) 0.10

(c) 0.12

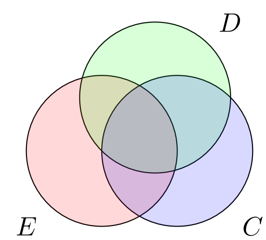
(d) 0.18

(e) None of the above.

Solution: (b)

Let E denote the event that the person chosen exercises regularly, D that he/she owns a dog and C that he/she likes cauliflower. The problem states that

$$\begin{aligned}\mathbb{P}[E|D] &= 0.6, \qquad \mathbb{P}[C \cap D] = 0.1, \qquad \mathbb{P}[C \cap D \cap E] = 0.04,\\ \mathbb{P}[E] &= 0.4, \qquad \mathbb{P}[D] = 0.3, \qquad \mathbb{P}[C] = 0.2.\end{aligned}$$



We are looking for $\mathbb{P}[E \cap C \cap D | C \cup D] = \mathbb{P}[E \cap C \cap D]/\mathbb{P}[C \cup D]$. We know that $\mathbb{P}[C \cup D] = \mathbb{P}[C] + \mathbb{P}[D] - \mathbb{P}[C \cap D] = 0.2 + 0.3 - 0.1 = 0.4$, so the required probability is $0.04/0.4 = 0.1$.

Problem 1.17. (5 points) The local pool supply store is having an end-of-season sale. They have a seemingly infinite number of floaties lying around. They know that among those 1/4 are ancient floaties from seasons past (so, old) and that 3/4 are the last season's floaties (so, new). We know that 15% of old floaties leak, and that 5% of new floaties leak. When an order comes in, a floatie is chosen at random to fulfill the order. You are excited about the sale and you are the first one to show up at the door. You buy a floatie. You take it to the pool. It leaks. What's the probability that it was an old floatie?

(a) 0.50

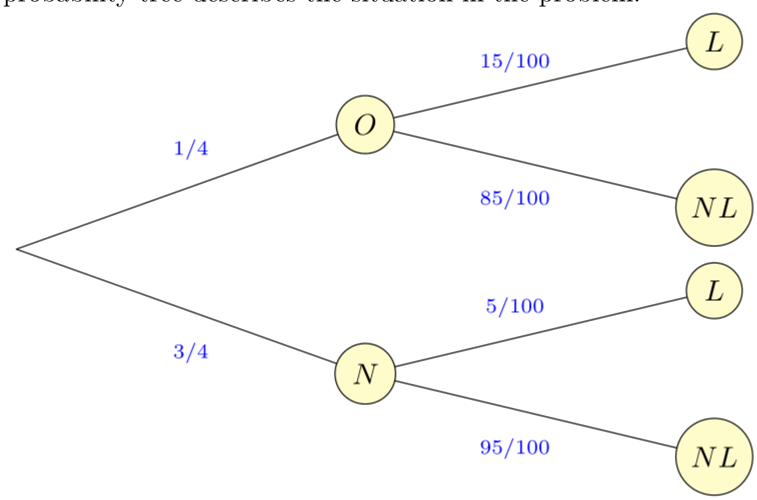
(b) 0.55

(c) 0.60

(d) 0.65

(e) None of the above.

Solution: (a)
This probability tree describes the situation in the problem:



We use the Bayes' theorem here.

$$\mathbb{P}[O \mid \text{floatie leaks}] = \frac{\mathbb{P}[\text{floatie leaks} \mid O]\mathbb{P}[O]}{\mathbb{P}[\text{floatie leaks}]}.$$

Using our tree, we get

$$\mathbb{P}[\text{floatie leaks}] = 0.25(0.15) + 0.75(0.05) = 0.075.$$

So,

$$\mathbb{P}[O \mid \text{floatie leaks}] = \frac{0.25(0.15)}{0.075} = 0.5.$$

Problem 1.18. (5 points) The local pool supply store is having an end-of-season sale. They have a seemingly infinite number of floaties lying around. They know that among those 1/4 are ancient floaties from seasons past (so, old) and that 3/4 are the last season's floaties (so, new). We know that 15% of old floaties leak, and that 5% of new floaties leak. When an order comes in, a floatie is chosen at random to fulfill the order. You are excited about the sale and you are the first one to show up at the door. You buy a floatie. You take it to the pool. It does not leak. What's the probability that it was a new floatie?

(a) About 0.77

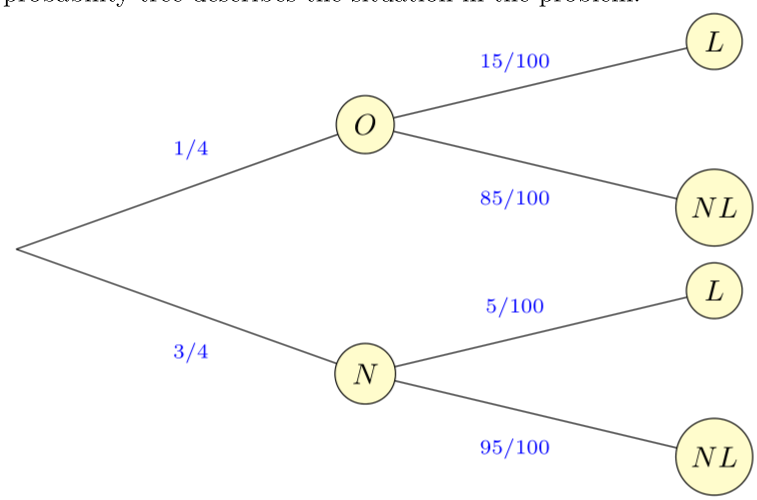
(b) About 0.79

(c) About 0.85

(d) About 0.87

(e) None of the above.

Solution: (a)
This probability tree describes the situation in the problem:



We use the Bayes' theorem here.

$$\mathbb{P}[N \mid \text{floatie does not leak}] = \frac{\mathbb{P}[\text{floatie does not leak} \mid N]\mathbb{P}[N]}{\mathbb{P}[\text{floatie does not leak}]}.$$

Using our tree, we get

$$\mathbb{P}[\text{floatie does not leak}] = 0.25(0.85) + 0.75(0.95) = 0.925.$$

So,

$$\mathbb{P}[N \mid \text{floatie does not leak}] = \frac{0.75(0.95)}{0.925} = 0.7703.$$

Problem 1.19. (5 points) Most mornings, Bertie Wooster asks Jeeves whether it is going to rain that day. It being England, Jeeves forecasts rain 80% of the time and dry weather the remaining 20% of the time. If Jeeves forecasts rain, the chance if it actually raining is 95%. If Jeeves forecasts no rain, the chance of it not raining is 90%. Suppose that one day Bertie forgot to ask Jeeves if it would rain. It did not rain. What is the probability that Jeeves would have predicted no rain?

(a) About 0.82

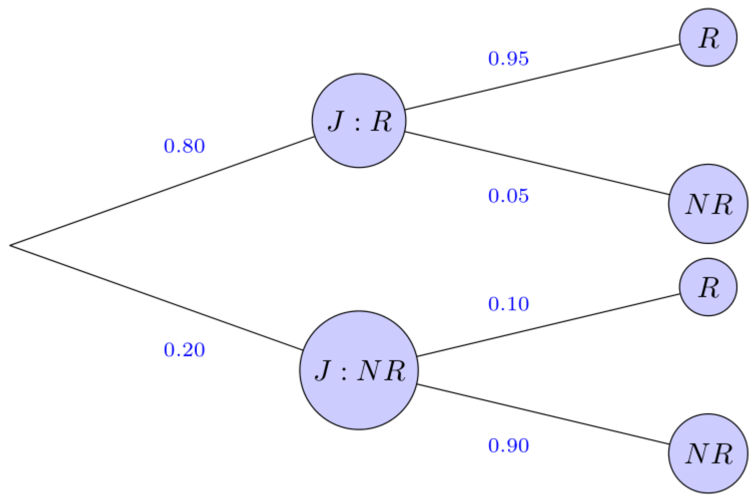
(b) About 0.84

(c) About 0.86

(d) About 0.88

(e) None of the above.

Solution: (a)
This probability tree describes the situation in the problem:



We use the Bayes' theorem here.

$$\mathbb{P}[\text{Jeeves would have said no rain} \,|\, NR] = \frac{\mathbb{P}[NR \,|\, \text{Jeeves would have said no rain}]\mathbb{P}[\text{Jeeves would have said no rain}]}{\mathbb{P}[NR]}.$$

Using our tree, we get

$$\mathbb{P}[NR] = 0.8(0.05) + 0.2(0.90) = 0.22.$$

So,

$$\mathbb{P}[\text{Jeeves would have said no rain} \,|\, NR] = \frac{0.2(0.9)}{0.22} = \frac{9}{11} = 0.82..$$

Problem 1.20. (5 points) Most mornings, Bertie Wooster asks Jeeves whether it is going to rain that day. It being England, Jeeves forecasts rain 80% of the time and dry weather the remaining 20% of the time. If Jeeves forecasts rain, the chance if it actually raining is 95%. If Jeeves forecasts no rain, the chance of it not raining is 90%. Suppose that one day Bertie forgot to ask Jeeves if it would rain. It rained. What is the probability that Jeeves would have predicted rain?

(a) About 0.9177

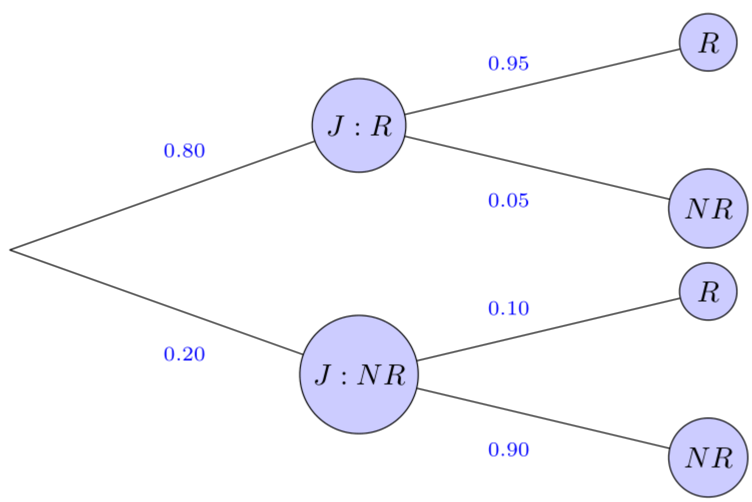
(b) About 0.9316

(c) About 0.9536

(d) About 0.9744

(e) None of the above.

Solution: (d)
This probability tree describes the situation in the problem:



We use the Bayes' theorem here.

$$\mathbb{P}[\text{Jeeves would have said rain} \,|\, R] = \frac{\mathbb{P}[R \,|\, \text{Jeeves would have said rain}]\mathbb{P}[\text{Jeeves would have said rain}]}{\mathbb{P}[R]}\,.$$

Using our tree, we get

$$\mathbb{P}[R] = 0.8(0.95) + 0.2(0.10) = 0.78.$$

So,

$$\mathbb{P}[\text{Jeeves would have said rain} \,|\, R] = \frac{0.8(0.95)}{0.78} = \frac{38}{39} = 0.974359\,.$$

Problem 1.21. (5 points) A box contains three coins; one coin is fair, one coin is two-headed and one coin is weighted so that the probability of heads is $1/3$. A coin is selected at random and tossed. You observe that the outcome of this coin-toss is tails. What is the probability that the randomly selected coin was the fair one?

(a) $2/7$

(b) $3/7$

(c) $4/7$

(d) $5/7$

(e) None of the above.

Solution: (b)
Let

$$E = \{\text{"the selected coin was fair"}\},$$
$$F = \{\text{"the selected coin was two-headed"}\},$$
$$G = \{\text{"the selected coin shows heads with probability } 1/3\text{"}\}.$$

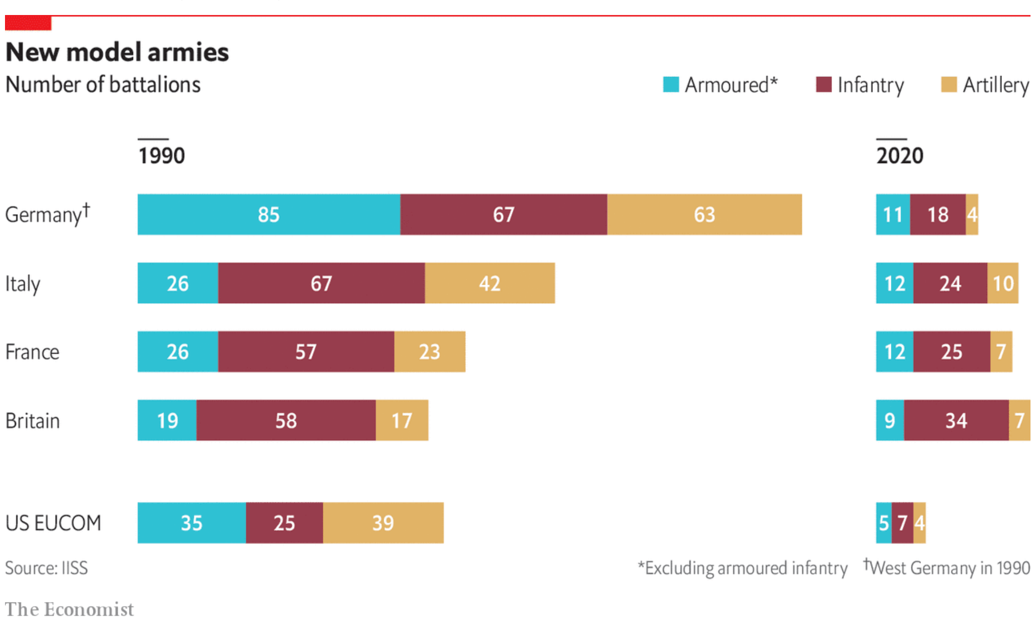
Define

$$B = \{\text{"the outcome of the coin-toss is tails"}\}.$$

We need to calculate $\mathbb{P}[E|B]$. Using Bayes' Theorem, we get that

$$\begin{aligned}\mathbb{P}[E|B] &= \frac{\mathbb{P}[E \cap B]}{\mathbb{P}[B]} \\ &= \frac{\mathbb{P}[B|E]\mathbb{P}[E]}{\mathbb{P}[B|E]\mathbb{P}[E] + \mathbb{P}[B|F]\mathbb{P}[F] + \mathbb{P}[B|G]\mathbb{P}[G]} \\ &= \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + \frac{2}{3}\cdot\frac{1}{3}} \\ &= \frac{\frac{1}{2}}{\frac{1}{2}+\frac{2}{3}} = \frac{3}{7}.\end{aligned}$$

Problem 1.22. (5 points) Consider the following charts and choose which of the offered statements is **correct**.



(a) European armies have overall increased in the last 30 years.

(b) The number of British **armoured** battalions is less than a third of what it used to be 30 years ago.

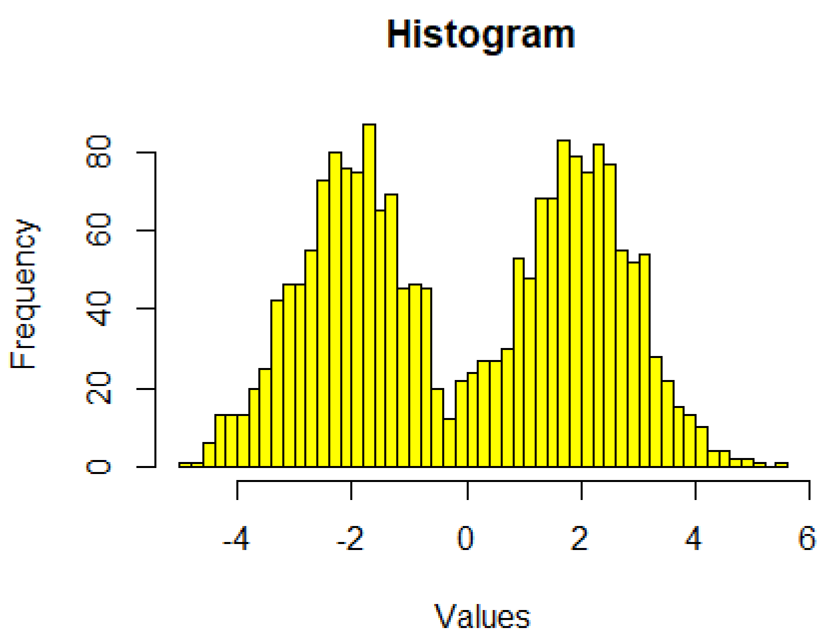
(c) The number of Italian battalions fell by 67%.

(d) French infantry troops fell by less than 50%.

(e) None of the above.

Solution: (e)
Since so many of you decided to round way up to 67%, I also accepted **(c)** as correct.

Problem 1.23. (5 points) Consider the following histogram:



The histogram is ...

(a) ...unimodal.

(b) ...bimodal, skewed.

(c) ...bimodal, symmetric.

(d) ...uniform.

(e) None of the above.

Solution: (c)

Problem 1.24. (5 points) *Source: AMC8, 2013.*

Hammie is in the 6^{th} grade and weighs 106 pounds. His quadruplet sisters are tiny babies and weigh 5, 5, 6, and 8 pounds. Which is greater, the average (mean) weight of these five children or the median weight, and by how many pounds?

(a) The median, by 60 pounds.

(b) The median, by 20 pounds.

(c) The mean, by 5 pounds.

(d) The mean, by 15 pounds.

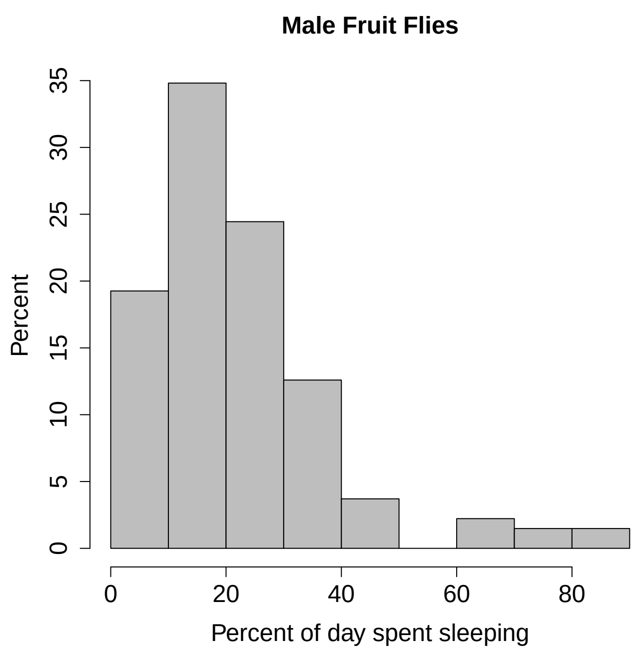
(e) The mean, by 20 pounds.

Solution: (e)

Lining up the weights in ascending order $(5, 5, 6, 8, 106)$, we see that the median weight is 6 pounds. The mean weight is

$$\frac{5+5+6+8+106}{5} = 26.$$

Problem 1.25. (5 points) The plot below displays the histogram of the percent of days spent sleeping by 100 male fruit flies. Which of the following is a sensible estimate of the mean and median of this distribution?



(a) mean = 24, median = 18

(b) mean = 18, median = 24

(c) mean = 18, median = 18

(d) mean = 20, median = 40

(e) mean = 25, median = 35

Solution: (a)

This is very similar to the textbook problem 2.12 (which was assigned as homework). Due to large values observed (a long tail), we realize that the distribution is right-skewed. So, the mean will be larger than the median. On the other hand, ”eyeballing” the heights of the rectangles in the histogram, we see that the median is a tad lower than 20. The only one of the offered pairs which qualifies is **(a)**.

Problem 1.26. (5 points) When a statistic, like the median, is said to be **robust**, this means that ...

(a) it is impossible for the data to have any outliers.
(b) the statistic is greatly influenced by the value of the outliers.
(c) the statistic is not greatly influenced by the value of the outliers.
(d) the statistic itself is an outlier.
(e) None of the above.

Solution: (c)

Problem 1.27. Below are some summary statistics from the `score` variable indicating employee satisfaction.

```
min Q1 median   Q3 max   mean       sd   n missing
 30 57   69.5   77  99 65.075 16.09361 200       0
```

Which of the following is true?

(a) The standard deviation estimate is not possible because `score` is a whole number.
(b) There is evidence that the distribution of `score` is right-skewed.
(c) The minimum value of 30 would be identified as an outlier in a box plot.
(d) There were more survey respondents who reported job satisfaction scores less than 57 than survey respondents who reported job satisfaction scores greater than 77.
(e) None of the above are true.

Solution: (e)

Problem 1.28. (5 points) Which of the following is true regarding the standard deviation?

(a) Standard deviation is a measure of central tendency.
(b) In the presence of skew, the standard deviation represents an observation's typical distance from the median rather than mean.
(c) It is possible to have a standard deviation of zero.
(d) It is possible to have a negative standard deviation.
(e) None of the above.

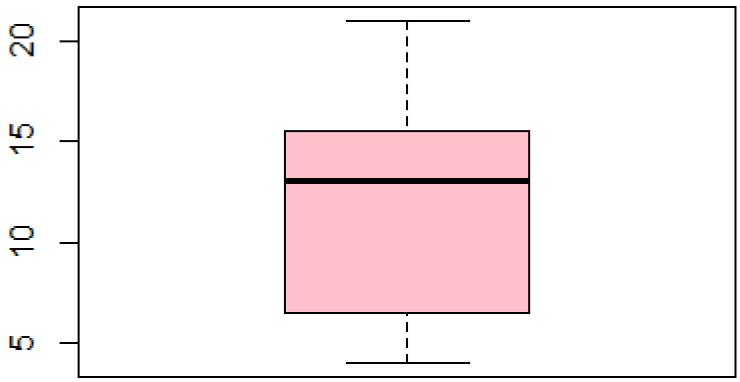
Solution: (c)

Problem 1.29. (5 points) A political scientist is interested in the effect of government type on economic development. She wants to use a sample of 30 countries evenly represented among the Americas, Europe, Asia, and Africa to conduct her analysis. What type of study should she use to ensure that countries are selected from each region of the world?

(a) Observational - simple random sample
(b) Observational - cluster sampling
(c) Observational - stratified sampling
(d) Observational - multiphase sampling
(e) None of the above.

Solution: (c)

Problem 1.30. (5 points) Consider the following box plot:



What do you suspect to be true about the data set?

(a) The distribution is symmetric.

(b) The maximum observation is 22.

(c) The minimum observation is 3

(d) The distribution is skewed.

(e) None of the above.

Solution: (b), (c), and (d)
Look at p. 49 from the book to see how the box plot is constructed and how to interpret it.

1.3. **Randomized response.**

Problem 1.31. (5 points) **Toddler Bribery!**

Dr. P. Piagette, a developmental psychologist, is trying to figure out what percentage of parents of small children resort to offering candy to their offspring to quiet them down. Realizing that parents might not answer truthfully to an outright question about bribing their little ones, she decides to use the randomized-response method.

She sets up a computer to display the question

"Have you ever offered candy to appease your toddler?"

with probability 0.6. The rest of the time, a virtual spinner spins on the screen. Half of the time, the spinner lands on red, a third of the time, the spinner lands on blue, and one sixth of the time, the spinner lands on yellow. The parent is asked

"Did the spinner land on red?"

In both cases, the subject is prompted to click the button with **Yes** or **No**. The interviewer did not know the actual question asked, just the ultimate response. So, there was no real reason for the subject to lie, and we assume that the subjects responded truthfully.

It turned out that 74% of the subjects answered "yes". Give an estimate of the proportion of *parents who bribe their children* in this population.

(a) 0.75
(b) 0.8
(c) 0.85
(d) 0.9
(e) None of the above.

Solution: (d)
Now, we are given that $\mathbb{P}[Yes] = 0.74$. Our goal is to figure out $p = \mathbb{P}[Yes \mid Q]$ with the conditioning event Q given by

$$Q = \{\text{the subject was asked the bribery question}\}.$$

We are given that $\mathbb{P}[Q] = 0.6$.

By the *Law of Total Probability*,

$$\begin{aligned}\mathbb{P}[Yes] = \mathbb{P}[Yes \cap Q] + \mathbb{P}[Yes \cap Q^c] &= \mathbb{P}[Yes \mid Q]\mathbb{P}[Q] + \mathbb{P}[Yes \mid Q^c]\mathbb{P}[Q^c] \\ &= p(0.6) + 0.5(0.4) = 0.6p + 0.2.\end{aligned}$$

So,

$$0.6p = 0.74 - 0.2 = 0.54 \quad \Rightarrow \quad p = 0.9.$$

Problem 1.32. (5 points) **Toddler Bribery!**

Dr. P. Piagette, a developmental psychologist, is trying to figure out what percentage of parents of small children resort to offering candy to their offspring to quiet them down. Realizing that parents might not answer truthfully to an outright question about bribing their little ones, she decides to use the randomized-response method.

She sets up a computer to display the question

"Have you ever offered candy to appease your toddler?"

with probability 0.6. The rest of the time, a virtual spinner spins on the screen. Half of the time, the spinner lands on red, a third of the time, the spinner lands on blue, and one sixth of the time, the spinner lands on yellow. The parent is asked

"Did the spinner land on blue?"

In both cases, the subject is prompted to click the button with **Yes** or **No**. The interviewer did not know the actual question asked, just the ultimate response. So, there was no real reason for the subject to lie, and we assume that the subjects responded truthfully.

It turned out that 53% of the subjects answered "yes". Give an estimate of the proportion of *parents who bribe their children* in this population.

(a) About 63%.

(b) About 66%.

(c) About 69%.

(d) About 72%.

(e) None of the above.

Solution: (b)
Now, we are given that $\mathbb{P}[Yes] = 0.53$. Our goal is to figure out $p = \mathbb{P}[Yes \mid Q]$ with the conditioning event Q given by

$$Q = \{\text{the subject was asked the bribery question}\}.$$

We are given that $\mathbb{P}[Q] = 0.6$.

By the *Law of Total Probability*,

$$\begin{aligned}\mathbb{P}[Yes] = \mathbb{P}[Yes \cap Q] + \mathbb{P}[Yes \cap Q^c] &= \mathbb{P}[Yes \mid Q]\mathbb{P}[Q] + \mathbb{P}[Yes \mid Q^c]\mathbb{P}[Q^c]\\ &= p(0.6) + \frac{1}{3}(0.4) = \frac{3}{5}p + \frac{2}{15}.\end{aligned}$$

So,

$$\frac{3}{5}p = 0.53 - \frac{2}{15} \quad \Rightarrow \quad p = 0.6611.$$

Problem 1.33. (5 points) **Toddler Bribery!**

Dr. P. Piagette, a developmental psychologist, is trying to figure out what percentage of parents of small children resort to offering candy to their offspring to quiet them down. Realizing that parents might not answer truthfully to an outright question about bribing their little ones, she decides to use the randomized-response method.

She sets up a computer to display the question

"Have you ever offered candy to appease your toddler?"

with probability 0.6. The rest of the time, a virtual spinner spins on the screen. Half of the time, the spinner lands on red, a third of the time, the spinner lands on blue, and one sixth of the time, the spinner lands on yellow. The parent is asked

"Did the spinner land on yellow?"

In both cases, the subject is prompted to click the button with **Yes** or **No**. The interviewer did not know the actual question asked, just the ultimate response. So, there was no real reason for the subject to lie, and we assume that the subjects responded truthfully.

It turned out that 52% of the subjects answered **Yes**. Give an estimate of the proportion of *parents who bribe their children* in this population.

(a) About 56%.

(b) About 66%.

(c) About 76%.

(d) About 86%.

(e) None of the above.

Solution: (c)

Now, we are given that $\mathbb{P}[Yes] = 0.52$. Our goal is to figure out $p = \mathbb{P}[Yes \mid Q]$ with the conditioning event Q given by

$$Q = \{\text{the subject was asked the bribery question}\}.$$

We are given that $\mathbb{P}[Q] = 0.6$.

By the *Law of Total Probability*,

$$\begin{aligned}\mathbb{P}[Yes] = \mathbb{P}[Yes \cap Q] + \mathbb{P}[Yes \cap Q^c] &= \mathbb{P}[Yes \mid Q]\mathbb{P}[Q] + \mathbb{P}[Yes \mid Q^c]\mathbb{P}[Q^c] \\ &= p(0.6) + \frac{1}{6}(0.4) = \frac{3}{5}p + \frac{1}{15}.\end{aligned}$$

So,

$$\frac{3}{5}p = 0.52 - \frac{1}{15} \quad \Rightarrow \quad p = 0.7556.$$