Workshop Notes

Dan Tran

2022-09-12

Contents

Workshop Recordings can be found on the Learning Resources Site: https://blackboard.qut.edu.au/bbcswebdav/pid-1481644-dt-announcement-rid-58665563_1/courses/MXB107_22se2/_site/videos.html.

In this document, I'll only write things that are not included in the official Workshop resources (above).

Workshop 4 (23/8/22)

Largest Dice Roll

Imagine everyone in this class is lost in a far-away desert. There is not enough food, so we decide who get to eat by rolling two dices.

Team A would win if the largest face rolled between the two is 1,2,3 or 4.

Team B win otherwise (if the largest face is 5 or 6).

Let A be the event that team A win, and B the event that team B win.

	1	2	3	4	5	6
1	A	A	A	A	B	B
2	A	A	A	A	B	B

	1	2	3	4	5	6
3	A	A	A	A	B	B
4	A	A	A	A	B	B
5	B	B	B	B	B	B
6	B	B	B	B	B	B

By counting, we observe that:

$$Pr(A) = \frac{16}{36} \approx 44.4\%$$
$$Pr(B) = \frac{20}{36} \approx 55.6\%$$

We see that team B actually wins more often than team A.

Bayesian: Rare Disease

There is a rare yet deadly disease that affects 0.1% of the population. To diagnose the disease, scientists have made a test kit with 99% accuracy. This is, if the patient indeed has the disease, it will test positive 99% of the time (and if you are sick, there is a 1% chance that it says negative). This also means that even if you are not sick, it can give positive result 1% of the time.

Positive Once What is the probability that you have the disease given that you tested positive once?

Answer

Let S be the event that you are sick, and P be the event that you test positive.

$$\begin{aligned} Pr(S) &= 0.001 \\ Pr(P|S) &= 0.99 \\ Pr(P|S^c) &= 0.01 \\ Pr(P^c|S) &= 0.01 \end{aligned}$$

We need to find $Pr(S|P)$, Bayesian Theorem would state that:

$$\begin{aligned} Pr(S|P) &= \frac{Pr(P|S)Pr(S)}{Pr(P)} \\ Pr(P) &= Pr(P|S)Pr(S) + Pr(P|S^c)Pr(S^c) \quad \text{(Law of total probability)} \end{aligned}$$

Substitution:

$$\begin{aligned} \Rightarrow Pr(S|P) &= \frac{Pr(P|S)Pr(S)}{Pr(P|S)Pr(S) + Pr(P|S^c)Pr(S^c)} \\ &= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.01 \times 0.999} \\ &\approx 0.0901 \approx 9.01\% \end{aligned}$$

Hence, if you tested positive once, the probability that you actually have the disease is 9.01% (which is not that bad!).

Positive Twice* *This is a challenging question for extension.

Now, you took another test and got a positive result. What is the probability that you have the disease now?

Answer

Renamed the above P to P_1, we have:

$$Pr(S|P_1) \approx 0.0901$$

For two tests, we have:

$$Pr(S|P_2 \cap P_1) = \frac{Pr(P_2 \cap P_1|S)Pr(S)}{Pr(P2 \cap P_1|S)Pr(S) + Pr(P_2 \cap P_1|S^c)Pr(S^c)}$$

Realise that P_1 and P_2 are independent, we see that

$$Pr(P_2 \cap P_1|S) = Pr(P_2|S)Pr(P_1|S)Pr(P_2 \cap P_2|S^c) = Pr(P_2|S^c)Pr(P_1|S^c)$$

Hence,

$$\begin{aligned}
Pr(S|P_2 \cap P_1) &= \frac{Pr(P_2 \cap P_1|S)Pr(S)}{Pr(P2 \cap P_1|S)Pr(S) + Pr(P_2 \cap P_1|S^c)Pr(S^c)} \\
&= \frac{Pr(P_2|S)Pr(P_1|S)Pr(S)}{Pr(P_2|S)Pr(P_1|S)Pr(S) + Pr(P_2|S^c)Pr(P_1|S^c)Pr(S^c)} \\
&= \frac{0.99 \times 0.99 \times 0.001}{0.99 \times 0.99 \times 0.001 + 0.01 \times 0.01 \times 0.99} \\
&\approx 0.9082 \approx 90.82\%
\end{aligned}$$

So the new probability is now 90.82%.

Workshop 5 (30/8/22)

Sum of Two Dices

Let X be the random variable generated by summing the results of two dices.

a. Find the probability mass function

Table 2: Remember, probability mass function is the probability of each event X:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$p(x) = Pr(X = x), x \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

By counting, we find that

$$p(x) = \begin{cases} \frac{1}{36} & , x = 2, 12 \\ \frac{2}{36} & , x = 3, 11 \\ \frac{3}{36} & , x = 4, 10 \\ \frac{4}{36} & , x = 5, 9 \\ \frac{5}{36} & , x = 6, 8 \\ \frac{6}{36} & , x = 7 \end{cases}$$

b. Find the mean

$$\begin{aligned} E[X] &= \sum_S xp(x) \\ &= \sum_{x=2}^{12} xp(x) \\ &= \frac{1}{36} \times 2 + \frac{2}{36} \times 3 + \frac{3}{36} \times 4 + \frac{4}{36} \times 5 + \frac{5}{36} \times 6 + \frac{6}{36} \times 7 \\ &+ \frac{5}{36} \times 8 + \frac{4}{36} \times 9 + \frac{3}{36} \times 10 + \frac{2}{36} \times 11 + \frac{1}{36} \times 12 \\ &= 7 \end{aligned}$$

Hence the mean is 7.

c. Find the median

$$\text{median of X} = \left\{ x \in S | Pr(X \leq x) \geq \frac{1}{2}, Pr(X \geq x) \geq \frac{1}{2} \right\}$$

Try $x = 8$:

$$\begin{aligned} Pr(X \leq 8) &= Pr(X = 1) + Pr(X = 2) + Pr(X = 3) + \cdots + Pr(X = 8) \\ &= \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} \\ &\approx 0.72 \geq \frac{1}{2} \\ Pr(X \geq 8) &= Pr(X = 8) + Pr(X = 9) + Pr(X = 10) + Pr(X = 11) + Pr(X = 12) \\ &= \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} \\ &\approx 0.42 \leq \frac{1}{2} \end{aligned}$$

Hence $x = 8$ is not the median of X.

Try $x = 7$:

$$
\begin{aligned}
Pr(X \leq 7) &= Pr(X=1) + Pr(X=2) + Pr(X=3) + \cdots + Pr(X=8) \\
&= \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} \\
&\approx 0.58 \geq \frac{1}{2} \\
Pr(X \geq 7) &= Pr(X=8) + Pr(X=9) + Pr(X=10) + Pr(X=11) + Pr(X=12) \\
&= \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} \\
&\approx 0.58 \geq \frac{1}{2}
\end{aligned}
$$

Hence $x = 7$ is the median of X.

d. Find the mode

$$\text{mode of X} = \max_{x \in S} p(x)$$

In this case, we know

$$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

and based on the $p(x)$ defined above, the most common sum is 7, where $p(7) = \frac{6}{36}$.

Hence the mode is 7.

Since mean = median = mode, we say X is defined by a symmetric, uni-modal distribution (i.e. no skew).

e. Find the variance and standard deviation

$$
\begin{aligned}
Var[X] &= E[(X - \mu]^2] \\
&= \sum_S (x - \mu)^2 p(x) \\
&= \sum_{x=2}^{12} (x - 7)^2 p(x) \\
&= (2-7)^2 \times \frac{1}{36} + (3-7)^2 \times \frac{2}{36} + (4-7)^2 \times \frac{3}{36} + \cdots + (11-7)^2 \times \frac{2}{36} + (12-7)^2 \times \frac{1}{36} \\
&\approx 5.83 \\
\sigma &\approx \sqrt{5.83} \approx 2.41
\end{aligned}
$$

Hence the variance of X is 5.83 and the standard deviation is 2.41.

Workshop 6 (6/9/22)

Workshop 7 (13/9/22)

Workshop 8 (20/9/22)

Workshop 9 (4/10/22)

Workshop 10 (11/10/22)

Workshop 11 (18/10/22)

Workshop 12 (23/10/22)