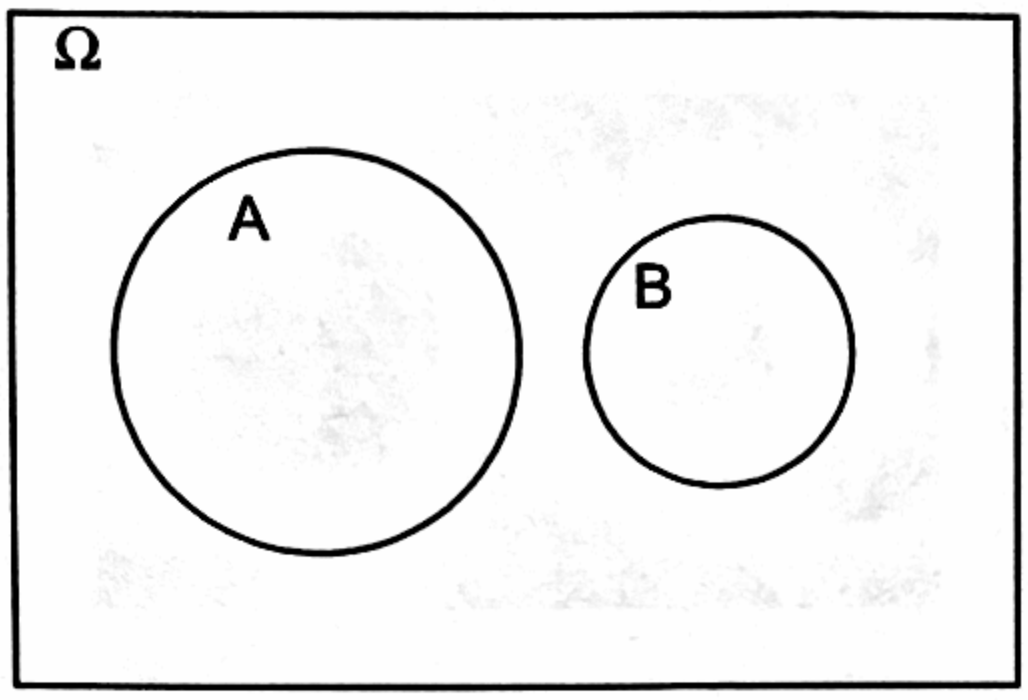
13.

Are the events A and B shown in the picture independent? Carefully justify your answer in each case.

a) [2 points]



A and B disjoint.

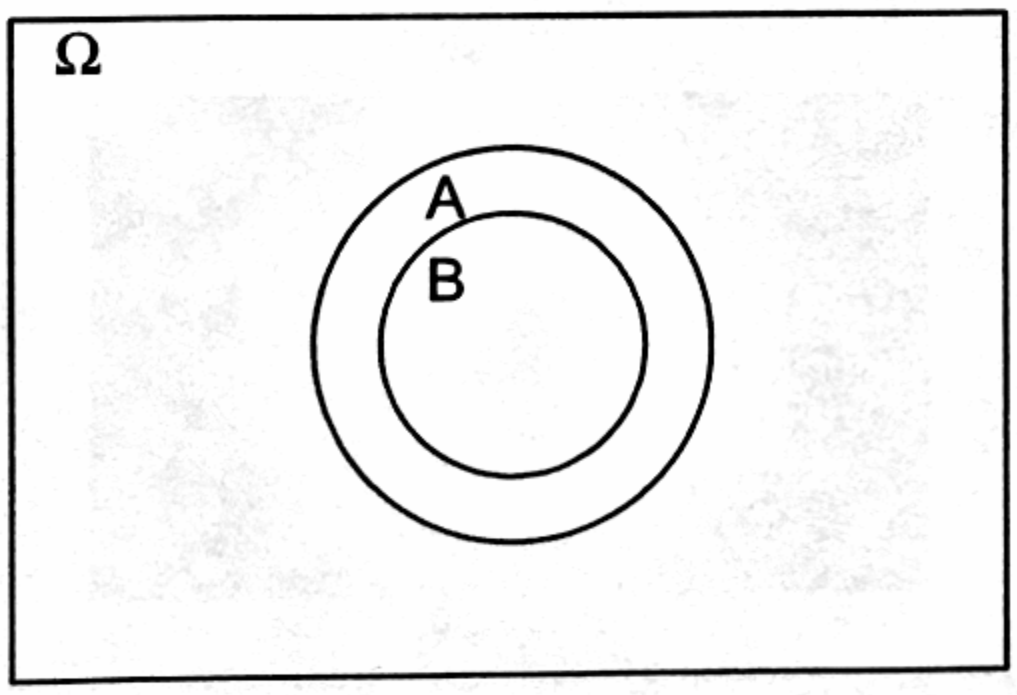
$P(A \cap B) = 0$.

But $P(A) \neq 0$, $P(B) \neq 0$.

So, $P(A \cap B) \neq P(A)P(B)$.

A and B are NOT independent.

b) [2 points]



Whenever the event B happens, A also happens.

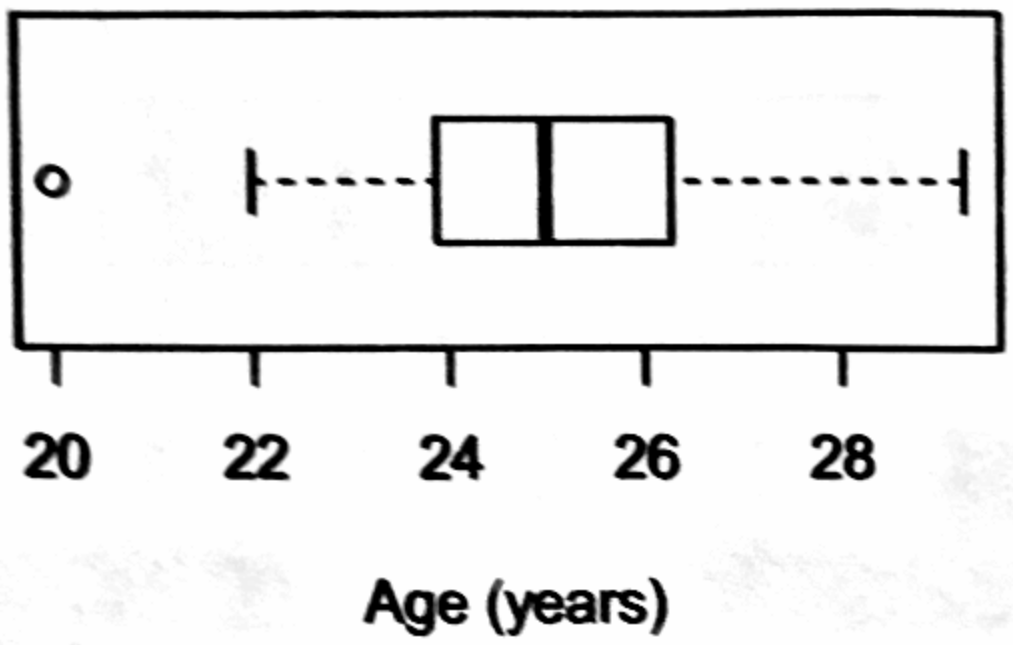
So, $P(A|B) = 1$.

But $P(A) \neq 1$.

So, $P(A|B) \neq P(A)$.

A and B are NOT independent.

14. Here is a box plot of ages of students in a certain class.



a) [2 points] Approximately 75% of students are at least <u>24</u> years old.

b) [2 points] Approximately <u>50</u>% of students are at least 25 years old.

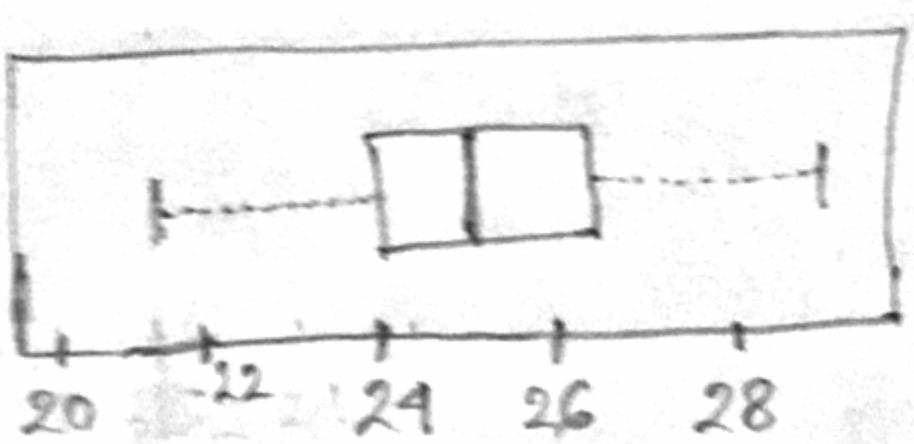
b) [2 points] Explain why the lower whisker does not extend to the 20 year old student.

$Q_1 = 24$., $IQR = Q_3 - Q_1 = 26 - 24 = 2$.

$Q_1 - 1.5 \times IQR = 24 - 1.5 \cdot 2 = 21$. (max. reach of lower whisker).

Since 20 is lower than this reach, the whisker does not extend to this student.

b) [2 points] Suppose the 20 year old drops the class but a 21.5 year old joins the class. Draw what the boxplot would look after these changes.



16. [4 points]

Suppose that $P(A|B) = P(A)$. Does this imply that $P(B|A) = P(B)$?

$$\frac{P(A\cap B)}{P(B)} = P(A)$$

$$\Rightarrow P(A\cap B) = P(A)P(B)$$

So, $P(B|A) = \dfrac{P(B\cap A)}{P(A)} = \dfrac{P(A\cap B)}{P(A)} = \dfrac{P(A)P(B)}{P(A)} = P(B).$

17. The fire alarm in your apartment building will go off (nearly) 100% of the time when there is a fire. However, the alarm is not perfect, especially during the summer. On a very hot and humid day, it has a 5% chance of going off even if there is no fire. The fire department knows that there have been only 2 real fires in the last 1000 hot and humid days and has no reason to think that there has been a rise or drop in the rate of fires.

a) [1 point] Based on this last sentence, what is the approximate probability that there is an actual fire on a given hot and humid day?

Let F : event that there is an actual fire on a given hot and humid day.

$P(F) = 2/1000 = 0{\cdot}002.$

b) [2 points] On what fraction of hot and humid days do we expect the alarm to go off?

Let A : alarm goes off.

Given $P(A|F) = 1$, $P(A|F^c) = 0{\cdot}05$.

From part (a), $P(F^c) = 1 - P(F) = 1 - 0{\cdot}002$.

$$P(A) = P(F)\,P(A|F) + P(F^c)\,P(A|F^c)$$
$$= 0{\cdot}002 \times 1 + (1 - 0{\cdot}002) \times 0{\cdot}05$$

c) [2 points] Suppose the alarm in your building goes off on a hot and humid day. What is the probability that there is a fire?

$$P(F|A) = \frac{P(F \cap A)}{P(A)} = \frac{P(F)\,P(A|F)}{P(F)\,P(A|F) + P(F^c)\,P(A|F^c)}$$
$$= \frac{0{\cdot}002 \times 1}{0{\cdot}002 \times 1 + (1 - 0{\cdot}002) \times 0{\cdot}05}$$

d) [4 points] What is the sensitivity and specificity of the alarm? (Hint: Think of a fire as having a disease, and think of the alarm going off as a test giving a positive.)

Sensitivity $= P(A|F) = 1$

Specificity $= P(A^c|F^c) = 1 - P(A|F^c)$

$= 1 - 0{\cdot}05.$

18. You work at the dairy bar, serving ice cream, on Friday afternoons. You know that on average 4 people buy an ice cream every minute on Friday afternoons. Let X be the number of people who will buy an ice cream between 2:00pm and 2:01pm next Friday afternoon.

a) [3 points] What sort of probability distribution would be appropriate for modeling X? Justify your answer. In addition to naming the type of distribution, also specify the values of all parameters.

A Poisson distribution will be appropriate, since
(i) customers buy ice creams at random time points independent of each other;
(ii) There is no upper bound on the number of customers who will buy ice cream between 2:00pm and 2:01pm.

$X \sim \text{Poisson}(\lambda)$, with $\lambda = 4$.

b) [2 points] What is the chance that no one buys an ice cream in that minute?

$$P(X=0) = e^{-4} \cdot \frac{4^0}{0!}.$$

c) [2 points] What is the chance that more than 2 people buy ice cream in that minute?

$$\begin{aligned} P(X>2) &= 1 - P(X \leq 2) \\ &= 1 - [P(X=0) + P(X=1) + P(X=2)] \\ &= 1 - \left[e^{-4} \cdot \frac{4^0}{0!} + e^{-4} \cdot \frac{4^1}{1!} + e^{-4} \cdot \frac{4^2}{2!}\right]. \end{aligned}$$

d) [3 points] What is the expected value and standard deviation of the number of people buying ice cream in that minute?

$E(X) = \lambda = 4$

$Var(X) = \lambda = 4$, so

$SD(X) = \sqrt{Var(X)} = \sqrt{\lambda} = \sqrt{4} = 2.$

19. Suppose you are taking a multiple choice exam with 10 questions and 4 choices per question. Suppose you are random guessing (you don't read the question, you just select a random choice on each question). Let X be the number of correct answers you get.

a) [2 points] What is the distribution of the random variable X? (Be sure to give the name of the distribution and the values of its parameters.)

$X \sim \text{Binomial}(n, p)$ with $n = 10$, $p = 1/4$

b) [2 points] What is the probability of getting a 90% or higher on the exam?

$$P(X \geq 9) = P(X=9) + P(X=10)$$
$$= \binom{10}{9} (0.25)^9 (1-0.25)^1 + \binom{10}{10} (0.25)^{10} (1-0.25)^0$$

c) [4 points] Suppose instead that you have studied for this exam, so that on any given problem you have a 60% chance of getting the right answer. Your friend has studied less and has a 40% chance of getting the right answer on any given problem. We would like to know the probability that you get a higher score on the exam than your friend. Describe (in words – no need to write any R code!) the method you could use to answer this question using a computer. Give

i. the general name of the technique you would use and also
ii. describe the basic idea behind what you are doing (imagine you are describing this to someone in your research group who is not in our class and does not know R). You should include in your explanation the definition of probability.