Practice exam (3)

Q1 A medical research team wishes to assess the usefulness of a certain symptom(S) in diagnosing a particular disease (D). In a random sample 780 persons with the disease,760 had the symptom. In another independent random sample of 1380 persons without the disease, only 21 had the symptom. Given that the disease prevalence in the population is 0.001, compute.

  (a) the symptom's false positive and false negative rates?
  (b) the predictive value positive for this symptom?
  (c) the specificity of the symptom?

Q2 A class of 80 university first-year students taking Statistics course were distributed by their gender and subject field (i.e. major) as follows:

Gender\Major	Information Technology	Basic Science	Engineering	Health Science	Total
Males	12	10	9	14	45
Females	5	7	6	17	35
Total	17	17	15	31	80

  (a) What proportion (i.e. percentage) of students not majoring in Engineering?
  (b) What percentage of students who are females or not majoring in Engineering?
  (c) Select a student randomly, what is the probability this student is female and not majoring in Engineering?
  (d)) Select a student randomly, what is the probability this student is female if she is majoring in Engineering?
  (e) Are student major and student gender independent? Justify

Q3 suppose that 25% of pregnant women (event G) suffer from anemia symptoms (event A) whereas only 10% of nonpregnant women suffer anemia symptoms. Suppose that in a group of 50 women, 20 are pregnant . assuming we select from this group a women at random, compute:
  a) The probability that she is pregnant.
  b) The probability that she is pregnant and suffers anemia.
  c) The probability that she suffers anemia symptoms.
  d) The probability that she is pregnant if she suffers anemia symptoms.

Q4 The data set 1,2,3,3,a,7,7,7,8,9,11 is bimodal, then a=

  A) 3   B) 4   C) 5   D) 6   E) 7

Q5 A national study found that treating people appropriately for high blood pressure reduced their overall mortality by 20%. Treating people adequately for hypertension has been difficult because it is estimated that 50% of hypertensives do not know they have high blood pressure, 50% of those who do know are inadequately treated by their physicians, and 50% who are appropriately treated fail to follow this treatment by taking the right number of pills.

A) What is the probability that among 10 true hypertensives at least 50% are being treated appropriately and are complying with this treatment?

B) What is the probability that at least 7 of the 10 hypertensives know they have high blood pressure?

Q6 Which of the following describes a parameter?

(A) A survey of 500 adults in Zarqa shows that 55% of them use public transportation

(B) A study of a sample of 2000 students in a university shows that 65% of them have done voluntary work

(C) A study of a sample of 3000 adults in Salat shows that 18% of them are Unemployed

(D) A survey of 40 employees in a company shows that 50% of them work overtime

(E) The average age at which all the presidents of a country were elected is 55 years

Q7 Which of the following statements is false?
(A) Any data entry greater than Q_1 -1.5(IQR) is an outlier
(B) The upper decile is P_{90} while the lower decile is P_{10}.
(C) When a data set is approximately symmetric, the mean equals the median.
(E) The third quartile is P75 of an ordered data set.

Q8 In a survey, 70% of college students said they drink coffee in the morning. Of those students who drink coffee in the morning, 20% get up late and of those who do not drink coffee in the morning, 30% get up late. A randomly selected college student is getting up late, find the probability that this student will drink coffee in the morning.

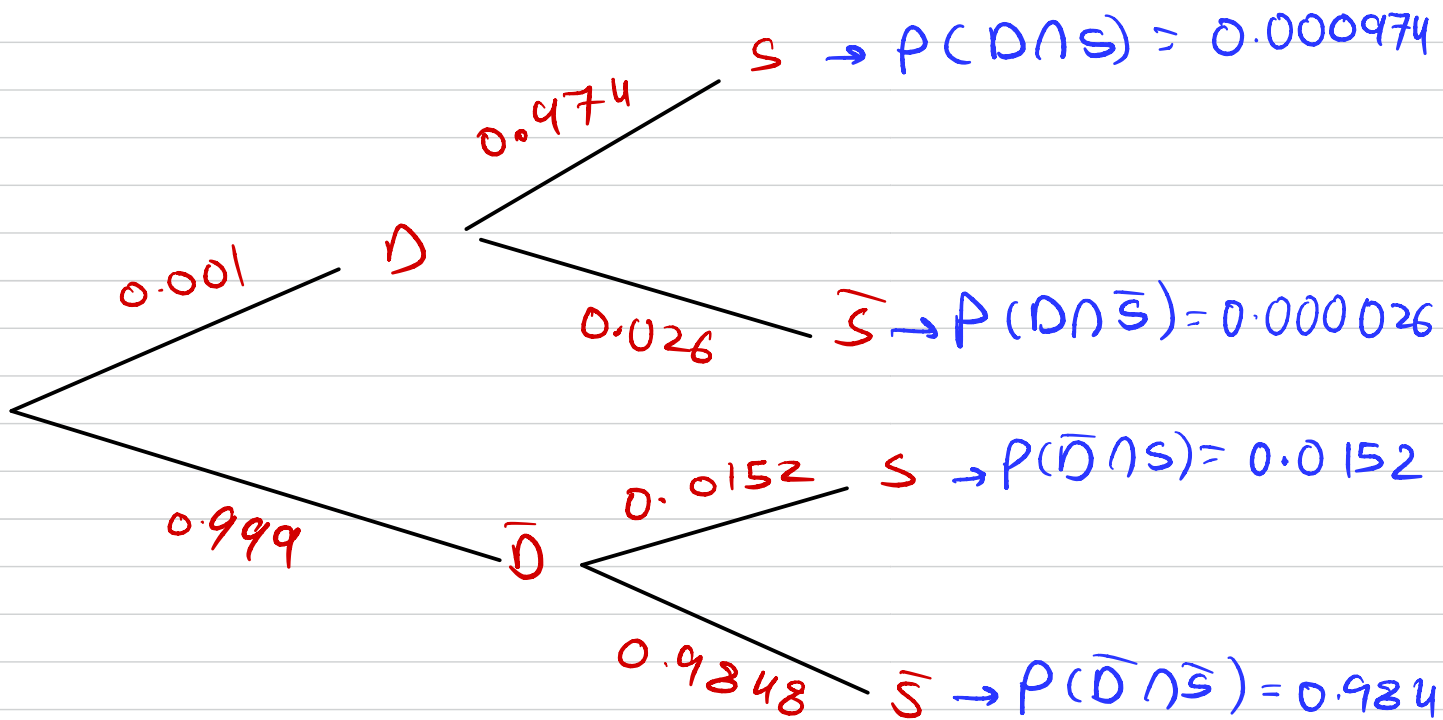(A) 4/13   (B) 9/13   (C) 2/9   (D) 7/9   (E) 14/23

Answers :-

Q_1 $P(S|D) = \dfrac{760}{780}$, $P(S|\bar{D}) = \dfrac{21}{1381}$, $P(D) = 0.001$

Tree diagram:

- D branch (prob 0.001):
  - S (prob 0.974) → $P(D \cap S) = 0.000974$
  - $\bar{S}$ (prob 0.026) → $P(D \cap \bar{S}) = 0.000026$
- $\bar{D}$ branch (prob 0.999):
  - S (prob 0.0152) → $P(\bar{D} \cap S) = 0.0152$
  - $\bar{S}$ (prob 0.9848) → $P(\bar{D} \cap \bar{S}) = 0.984$

a)

False positive $= P(S|\bar{D}) = 0.0152$

False Negative $= P(\bar{S}|D) = 0.026.$

b) $P(PV+) = P(D|S) = \dfrac{P(D \cap S)}{P(S)}$

$$= \dfrac{0.000974}{0.000974 + 0.0152} = 0.061$$

c) specificity $P(\bar{S}|\bar{D}) = \dfrac{P(\bar{S} \cap \bar{D})}{P(\bar{D})}$

$$= \dfrac{0.984}{0.999} = 0.985$$

Q_2

Gender\Major	Information Technology	Basic Science	Engineering	Health Science	Total
Males	12	10	9	14	45
Females	5	7	6	17	35
Total	17	17	15	31	80

a) $P(\bar{E}) = 1 - P(E) = 1 - \dfrac{15}{80} = 0.8125.$

b) $P(F \cup \bar{E}) = P(F) + P(\bar{E}) - P(F \cap \bar{E})$

$$= \frac{35}{80} + \frac{65}{80} - \frac{29}{80} = 0.8875$$

$P(F \cap \bar{E}) = P(F) - P(F \cap E) = \dfrac{35}{80} - \dfrac{6}{80} = \dfrac{29}{80}$

c) $P(F \cap \bar{E}) = \dfrac{29}{80} = 0.3625.$

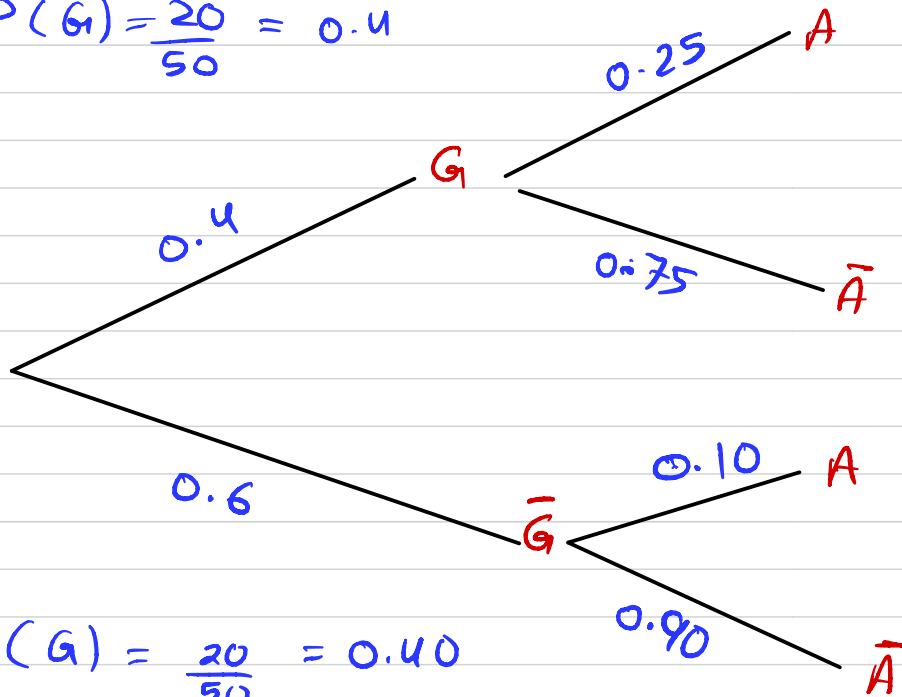d) $P(F \mid E) = \dfrac{P(F \cap E)}{P(E)} = \dfrac{6}{15} = 0.4$

e) lets check.

$P(F \mid E) = 0.4 \neq P(F) = 0.4375.$

Not independent. since we know.
$P(A \mid B) = P(A)$ if they are independent.

Q_3 $P(G) = \dfrac{20}{50} = 0.4$

Tree diagram:
- G branch: 0.4
  - A: 0.25
  - $\bar{A}$: 0.75
- $\bar{G}$ branch: 0.6
  - A: 0.10
  - $\bar{A}$: 0.90

a) $P(G) = \dfrac{20}{50} = 0.40$
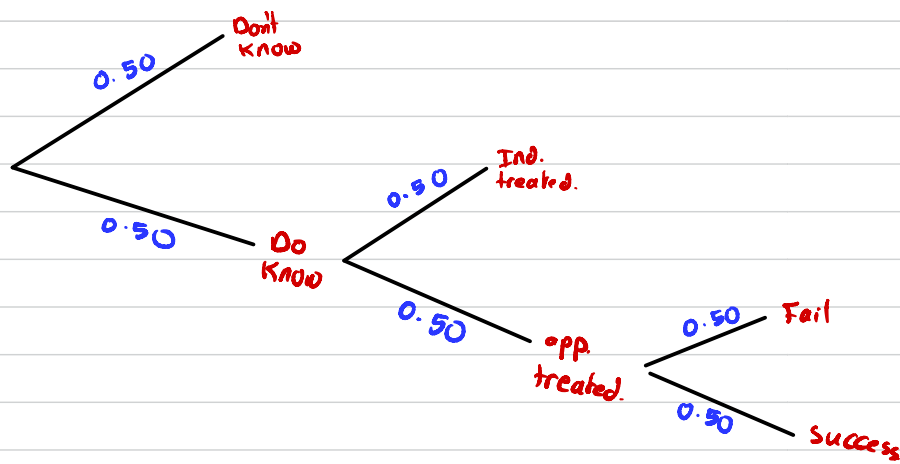
b) $P(G \cap A) = 0.40 \times 0.25 = 0.1$

c) $P(A) = P(G \cap A) + P(\bar{G} \cap A)$

$\qquad = 0.1 + 0.6 \times 0.10 = 0.16.$

d) $P(G|A) = \dfrac{P(G \cap A)}{P(A)} = \dfrac{0.1}{0.16} = 0.625$

Q_4 since it's bimodal $\Rightarrow$ $a = 3 \rightarrow$ Ⓐ

Q5

a) $P(\text{Do know} \cap \text{app. treated} \cap \text{fail}) = 0.5 \times 0.5 \times 0.5$

$$= 0.125.$$

$n = 10 \implies X \sim \text{Bin}(10, 0.125)$

at least 50% of 10 $\implies P(X \geqslant 5)$

$P(X \geqslant 5) = 0.00445.$

b) $n = 10$ but $P(\text{Do Know}) = 0.50.$

$X \sim \text{Bin}(10, 0.50)$

$P(X \geqslant 7) = 0.17188.$

Q6 answer is E.

Q7 answer is A.

Q 8.

$$P(C) = 0.70 \quad C$$

Tree diagram:

- 0.70 → C
  - 0.20 → h
  - 0.80 → $\bar{h}$
- 0.30 → $\bar{C}$
  - 0.30 → h
  - 0.70 → $\bar{h}$
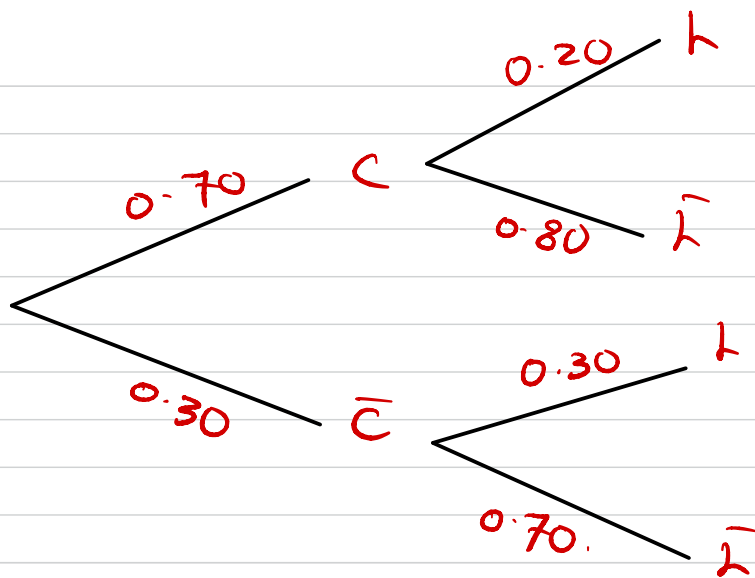
$$P(C \mid h) = \frac{P(C \cap h)}{P(h)}$$

$$= \frac{0.70 \times 0.20}{0.70 \times 0.20 + 0.30 \times 0.30} = \frac{14}{23}$$

(E)