EXAM OF STATISTICS (PROBABILITY AND RANDOM VARIABLES)

2nd Physiotherapy Version A **April, 27 2023**

Duration: 1 hour.

1. A water source contaminated contains 0.1 amoebas per litre on average.

   (a) What is the probability that 2 litres of water from this source contains more than one amoeba?

   (b) If 5 persons drink 2 litres of water from this source, what is the probability of having some person infected with amoebas?

   (c) If 100 persons drink half a litre of water from this source, what is the probability that less than 5 are infected with amoebas?

Solution

(a) Let X be the number of amoebas in 2 litres of contaminated water. Then $X \sim P(0.2)$ and $P(X > 1) = 0.0175$.

(b) The probability that a person who drank 2 litres of contaminated water is infected is $P(X \geq 1) = 0.1813$. Let Y be the number of persons infected with amoebas in a sample of 5 persons who drank 2 litres of contaminated water. Then $Y \sim B(5, 0.1813)$ and $P(Y \geq 1) = 0.6321$.

(c) Let U be the number of amoebas in half a litre of contaminated water. Then $U \sim P(0.05)$ and $P(U \geq 1) = 0.0488$. Let V be the number of persons infected with amoebas in a sample of 100 persons who drank half a litre of contaminated water. Then $V \sim B(100, 0.0488) \approx P(4.8771)$ and $P(V < 5) = 0.4623$.

---

2. Respiratory allergies affect 1 out of every 15 individuals in a population, while food intolerances affect 5% of individuals. Assuming that the two problems are independent,

   (a) Compute the probability of having at least one of the problems.

   (b) Compute the probability of having an allergy but not an intolerance.

   (c) Compute the probability of having neither of the two problems.

   (d) Compute the probability of having an allergy if you have an intolerance.

Solution

Let A the event of having respiratory allergies and B the event of having food intolerance.

(a) $P(A \cup B) = 0.1133$.

(b) $P(A - B) = 0.0633$.

(c) $P(\overline{A} \cap \overline{B}) = 0.8867$.

(d) $P(A|B) = 0.0667$.

---

3. In a population of 20000 women, it is known that back width follows a normal distribution with mean 29 cm and standard deviation 2.4 cm.

(a) Compute the number of women with a back width greater than 32 cm.

(b) Compute the interquartile range of women's back width and interpret it.

(c) Compute the probability that a woman with a back width above the third quartile, has a back width above 32.

Solution

Let X be the back width, then $X \sim N(29, 2.4)$.

(a) $P(X > 32) = 0.1056$ and approximately 2113 persons have a back width greater than 32 cm.

(b) $Q_1 = 27.3812$ cm, $Q_3 = 30.6188$ cm, and $IQR = 3.2376$ cm.

(c) $P(X > 32|X > 30.6188) = 0.4226$.

---

4. A diagnostic test for prostate cancer has a specificity of 80% and produces 1.6% of false negatives. It is known that the prevalence of prostate cancer in a population is 2%.

(a) Compute the sensitivity of the test. Does the outcome of the test depend on whether a man has prostate cancer?

(b) Is this a good test to diagnose the disease?

(c) What should be the minimum specificity of the test to diagnose the disease with a positive outcome?

Solution

(a) Sensitivity $= P(+|D) = 0.2$. The outcome of the test does not depend on the prostate cancer.

(b) Positive predictive value $= P(D|+) = 0.02 < 0.5$, so we can not confirm the prostate cancer with a positive outcome.

(c) Minimum specificity 0.9959.
