QUESTIONS IN PREPRATION FOR THE EXAMINATION

The examination asks for THREE answers

The time allowed is TWO hours

1. Establish what is meant by (a) a pair of mutually exclusive events and (b) a pair of statistically independent events.

Prove, with reference to the axioms of probability and the rules of boolean algebra, that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

The probability that a smoker who consumes over 40 cigarettes a day will suffer from chronic respiratory illness is 0.4. The probability that he will suffer from heart disease is 0.5. The probability that he will suffer from at least one of these ailments is 0.6

(i) Find the probability that the smoker will suffer from both ailments,

(ii) Find the probability that he will suffer from heart disease given that he suffers from chronic respiratory illness.

Answer. Let R denote respiratory illness and let H denote heart disease. We have

$$P(R) = \frac{4}{10}, \qquad P(H) = \frac{5}{10}, \qquad P(R \cup H) = \frac{6}{10}.$$

(i) The probability that the smoker will suffer from both ailments is

$$\begin{aligned} P(R \cap H) &= P(R) + P(H) - P(R \cup H) \\ &= \frac{4}{10} + \frac{5}{10} - \frac{6}{10} = \frac{3}{10}. \end{aligned}$$

(ii) The probability that he will suffer from heart disease given that he suffers from chronic respiratory illness is

$$P(H|R) = \frac{P(R \cap H)}{P(R)} = \frac{3}{10} . \frac{10}{4} = \frac{3}{4}.$$

2. Derive, from first principles, the function expressing the probability of obtaining x successes in n independent trials when the probability of a success in any trial is p.

On the third floor of the Metropolitan Hotel there are six guest rooms but only four bathrooms. On average, two guests in five require a morning bath. Calculate the probability that, on a morning when all the guest rooms are occupied, some of the bathrooms will have to be used more than once.

Answer: The probability that x out of 6 guests will take baths is given by

$$b\Big(x; n = 6, p = \frac{4}{10}\Big) = \frac{6!}{(6-x)!x!}\left(\frac{4}{10}\right)^x\left(\frac{6}{10}\right)^{6-x}.$$

The probability that bathrooms will have to be used more than once is

$$\begin{aligned} b(5) + b(6) &= 6\left(\frac{4}{10}\right)^5\left(\frac{6}{10}\right) + \left(\frac{4}{10}\right)^6 \\ &= \frac{4^5}{10^6}(36+4) = 0.040960. \end{aligned}$$

3. The probability of x arrivals in the time interval $[t, t+\tau]$ is given by the Poission function

$$p(x, \tau) = \frac{e^{-\lambda\tau}(\lambda\tau)^x}{x!}.$$

Show that

$$\sum_{x=0}^{\infty} p(x, \tau) = 1.$$

Also show

(a) that the probability of having to wait more that t periods for the first arrival is $P(w > t) = P(x = 0, t) = e^{-\lambda t}$,

(b) that the cumulative distribution of the waiting times is $F(t) = P(w < t) = 1 - e^{-\lambda t}$, and

(c) that the distribution of the arrival times is the function $f(t) = F'(t) = \lambda e^{-\lambda t}$

Given that $\lambda = 0.5$, what is the probability that you will have to wait for more than 4 minutes, given that you have already waited for 2 minutes?

Answer: The sum over x of the Poisson probabilities is

$$\sum_{x=0}^{\infty} p(x, t) = e^{-\lambda\tau}\sum_{x=0}^{\infty}\frac{(\lambda\tau)^x}{x!} = e^{-\lambda\tau} \times e^{\lambda\tau} = 1.$$

(a) The probability of having to wait more than t periods is obtained by setting $x = 0$ and $t = \tau$ in the Poisson formula above to give $P(w \geq t) = e^{-\lambda t}$.

(b) The cumulative distribution of the waiting times is $F(w) = P(w < t) = 1 - P(w \geq t) = 1 - e^{-\lambda t}$.

(c) The density function of the waiting times, which is the first derivate of the cumultive function, is $dF(t)/dt = f(t) = \lambda e^{-\lambda t}$.

The probability of having to wait for more than 4 minutes, given that one has already waited for two minutes is $P(w > 4|w > 2) = e^{-4\lambda}/e^{-2\lambda} = e^{-2\lambda} = e$

4. Demonstrate how the moments of a random variable x may be obtained from its moment generating function by showing that the rth derivative of $E(e^{xt})$ with respect to t gives the value of $E(x^r)$ at the point where $t = 0$.

Demonstrate that the moment generating function of a sum of independent variables is the product of their individual moment generating functions.

Find the moment generating function of the point binomial

$$f(x; p) = p^x(1 - p)^{1-x}$$

where $x = 0, 1$. What is the relationship between this and the m.g.f. of the binomial distribution?

Find the variance of x_1+x_2 when $x_1 \sim f(p_1 = 0.25)$ and $x_2 \sim f(p_2 = 0.75)$ are independent point binomials.

Answer. It is straightforward to show that $V(x_1) = V(x_2) = pq$ where $p = p_1$ and $q = p_2$ are used for ease of notation. Since x_1, x_2 are statistically independent, it follows that $V(x_1 + x_2) = V(x_1) + V(x_2) = 2pq$.

Alternatively, we may consider the moment generating functions of the two variables x_1 and x_2 which are respectively

$$M_1 = (p + qe^t) \qquad \text{and} \qquad M_2 = (q + pe^t).$$

The moment generating function of their sum $y = x_1 + x_2$ is

$$M = M_1M_2 = (p + qe^t)(q + pe^t) = pq + q^2e^t + p^2e^t + pqe^{2t}.$$

Its first and second derivatives are

$$\frac{dM}{dt} = q^2e^t + p^2e^t + 2pqe^{2t},$$

$$\frac{d^2M}{dt^2} = q^2e^t + p^2e^t + 4pqe^{2t};$$

and setting $t = 0$ gives the following moments:

$$E(y) = q^2 + p^2 + 2pq = 1,$$
$$E(y^2) = q^2 + p^2 + 4pq = 1 + 2pq.$$

It follows that

$$V(y) = E(y^2) - \{E(y^2)\} = 2pq = 0.375.$$

5. Let x and y be jointly distributed random variables such that $E(y|x) = \alpha + \beta x$. Show that $\beta = C(x, y)/V(x)$ and that $\alpha = E(y) - \beta E(x)$.

The centigrade temperatures recorded at 20 minute intervals in an air-conditioned room constitute a sequence of random variables. The expected value of the readings is 19° with a standard deviation of 2°. The correlation between successive temperature readings is 0.9. If a temperature of 20° is recorded at one reading, what is the expected value at the next reading?

Answer. We have

$$Corr\{y(t), y(t-1)\} = \frac{C\{y(t), y(t-1)\}}{V\{y(t)\}} = \beta = 0.9.$$

Therefore

$$\begin{aligned} E\{y(t+1)|y(t)\} &= \alpha + \beta y(t) \\ &= \{E(y_{t+1}) - \beta E(y_t)\} + \beta y_t \\ &= E(y_{t+1}) + \beta\{y_t - E(y_t)\} \\ &= 19 + 0.9[20 - 19] = 19.9. \end{aligned}$$

6. A factory that manufactures shafts of 5cm diameter has installed new lathes. Hitherto, the variance of the diameter of the shafts has been 0.49mm^2. A sample of 20 shafts, produced by the new machines, has a variance of 0.25mm^2 measured about the theoretical mean of 5cm. Find a 95% confidence interval for the new variance, and a 95% confidence interval for the ratio of the old and the new variances.

Answer. The previous variance was $\sigma_x^2 = 0.49$. The new variance is estimated as

$$\hat{\sigma}_y^2 = \frac{1}{n}\sum_{j=1}^{n}(y_j - \mu)^2 = 0.25, \qquad \text{where} \quad n = 20 \quad \text{and} \quad \mu = 5.$$

Moreover, there is

$$\frac{\sum_{j=1}^{n}(y_j - \mu)^2}{\sigma_y^2} = \frac{n\hat{\sigma}_y^2}{\sigma_y^2} \sim \chi^2(n).$$

Therefore, we can find numbers α and β from the tables of the $\chi^2(20)$ distribution such that

$$P\left(\alpha < \frac{n\hat{\sigma}_y^2}{\sigma_y^2} < \beta\right) = P\left(\frac{n\hat{\sigma}_y^2}{\beta} < \sigma_y^2 < \frac{n\hat{\sigma}_y^2}{\alpha}\right) = 0.95.$$

The values indicated by the tables are $\alpha = 10.851$ and $\beta = 31.410$. Therefore, the 95% confidence interval for the new variance is

$$\left[\frac{20 \times 0.25}{31.410}, \frac{20 \times 0.25}{10.851}\right] = [0.159, 0.461].$$

The confidence interval for the ratio of the old and the new variances is given by $[\sigma_x^2/0.159, \sigma_x^2/0.461]$ with $\sigma_x^2 = 0.49$.