BN001, BN009, BN012, BN117, BN121, BN903

Statistics and Probability

Worksheet – Random Variables

Question 1

A factory is producing components, of which 1.9% are defective. They are packed in boxes, each containing 15 components. Using the Poisson approximation to the Binomial distribution,

a) Calculate the probability that a box has 2 defective components.

b) It is estimated that a customer receiving a box of components will dismiss one defective but will complain if the box contains 2 or more defectives. Calculate the probability of this happening.

c) If 20 of these boxes are sold, identify the distribution for the random variable of the number of potential complaints.

Question 2

A multiple choice exam has 10 questions, each one with a choice of four answers, only one of which is correct. Let X be the random variable of the number of correct answers for a student picking their answers at random.

a) Identify the distribution governing variable X and therefore the expected value of correct questions for this student. Find the standard deviation for X.

b) Calculate the probability of the student getting no questions correct.

c) Calculate the probability of the student getting more than 3 questions correct, in other words a pass mark.

d) Consider the more general case of n questions, where each question has k possible answers with only one correct, so $pk = 1$. Create a graph of the function $f(n) = P[X > 0.4n]$ for $k = 3$ to 6, that is, the probability of passing against the number of question for a given number of answers. Describe what happens as n increases, explaining why. Repeat this for $k = 2$.

Question 3

A guest at a gambling club plays a game for 10€, where a player asks them to choose a number 1 to 6, then rolls three dice. If the chosen number does not come up, the guest does not win. If the number comes up once, the guest wins 10€, twice wins 20€ and three times wins 30€.

a) Identify the exact distribution for the number of times the number chosen by the guest comes up.

b) Using the expected value as a measure, calculate whether the game favours the guest or the player. Explain why the expected value is a good measure for this.

Question 4

Show that for a small value of p, the Binomial distribution becomes the Poisson as n becomes very large.

Question 5

A company operates a tollgate on an Autoroute in the Pyrénées. It has been established that for the month of October, cars arrive at this booth at a rate of one every 3 minutes.

a) Calculate the probability that a car arrives within 6 minutes.

b) Calculate the probability that 3 cars arrive in a ten-minute period for this tollbooth.

c) For another tollbooth, the rate is one car every 4 minutes. Calculate the probability that another car arrives within 8 minutes.

d) Compare the answers for part (a) and (b), explaining any similarities/differences.

Question 6

A communications device relies on a power source that experiences a power surge once every 4 days.

a) Calculate the probability that a power surge occurs within the next 2 days, and then for the next 6 days.

b) It is suspected that two power surges within one day will impair the functions of the device. Calculate the probability of this happening.

Question 7

Two players are in a game of Russian roulette. The revolver has six chambers and one live round is being used. After each play, the cylinder is rotated so that the next play is an independent event. The game ends once the revolver has fired and that player who fired the revolver loses. Let variable N be the number of shots fired in a game.

a) Use the laws of probability to write down a distribution for the variable N. Need the possibility of the game continuing endlessly be included?

b) Calculate the probability that the same player who starts the game loses. [Hint: recall that for any number k, $2k + 1$ is an odd number and $2k$ is an even number]

c) Identify the expected value and standard deviation for the variable N. [Hint: let the function f be given by

$$f(x) = a\sum x^k,$$

where the summation starts at $k = 0$. Write f in closed form and so use two parallel ways to calculate $f'(x)$. Repeat this process for $f''(x)$.]

d) Redo the calculation for part (b) for the case of two live rounds in the cylinder.

Question 7

The following questions use the normal approximation of the binomial distribution for large n.

a) Write down equations giving the values of the mean and standard deviation for this approximation in terms of the parameters n and p.

b) Assume that a pregnancy results in a boy or girl with equal probability. A large maternity hospital handles 75 births every week and a small one handles 25. Without actually calculating the probabilities, decide which hospital is more likely to produce a given proportion of girls above 0.5.

c) Calculate the probability that the proportion of girls exceeds a given value *p*, using the normal approximation. Comment on which of the probabilities in part (b) will be closest to this value.

Definition – The Binomial Distribution

A Bernoulli trial is being repeated, with a possible result *A*.

- Each time the trial is done, the probability of result *A* turning up, called a success, is *p*.
- The trial is repeated *n* times.

Let *X* be the random variable of the number of times event *A* comes up, the number of successes. The probability of getting *r* successes from *n* trials is:

$$P[X = r] = {}^{n}C_{r}\, p^{r}(1 - p)^{n - r}.$$

If *n* is large and *p* is small, so that *np* is not a large number, then this expression tends to the Poisson distribution below with $\mu = np$.

Definition – The Exponential Distribution

A series of incidents are occurring at a fixed rate of λ. Let T be the random variable of the time to the next incident. Then the variable follows the exponential distribution:

$$P[T < t] = 1 - e^{-\lambda t}.$$

Definition – The Poisson Distribution

A series of incidents are occurring at a fixed rate λ. The mean number of incidents happening per unit time is λ. Let T be a certain interval of time. Let N be the random variable of the number of these events which occur in a time T. Setting $\mu = \lambda T$, the random variable N has probability mass function:

$$P[N = r] = e^{-\mu}\frac{\mu^r}{r!}.$$

To approximate the Binomial Distribution $\mu = np$, as above.
