Ecarma, Destin R. - 23100411

1. Consider the following experiment:

   (a) The experiment consists of n identical trials.

   (b) The outcome of each trial falls into one of k classes or cells.

   (c) The probability that the outcome of a single trial falls into cell i, is $p_i, i = 1, 2, ..., k$ and remains the same from trial to trial. Notice that $p_1 + p_2 + p_3 + \ + p_k = 1.$

   (d) The trials are independent.

   (e) The random variables of interest are $Y_1, Y_2, ..., Y_k$, where Y_i equals the number of trials for which the outcome falls into cell i. Notice that $Y_1 + Y_2 + Y_3 + \ + Y_k = n$.

   (f) The joint probability distribution $P(y) = \frac{n!}{y_1!y_2!y_3!...y_n!} P_1^{y_1} P_2^{y_2} P_3^{y_3} ....P_n^{y_n}$

1.1 Show that the expected value $E[y] = \sum_y y \cdot P(y) = n \cdot p_i$

The expected value of a random variable is calculated as the sum of all possible values of the variable weighted by their respective probabilities.

For the random variable Y_i, representing the number of trials where the outcome falls into cell i, the possible values range from 0 to n.

$$E[Y_i] = \sum_{y_i=0}^{n} y_i \cdot P(Y_i = y_i) \tag{1}$$

Given the joint probability distribution function $P(y)$, where $P(y) = \frac{n!}{y_1!y_2!y_3!...y_n!} P_1^{y_1} P_2^{y_2} P_3^{y_3} ...P_n^{y_n}$, we can find $P(Y_i = y_i)$ by substituting y_i into the expression fro $P(y)$.

$$P(Y_i = y_i) = \frac{n!}{y_i!(n - y_i)!} p_i^{y_i} (1 - p_i)^{n-y_i} \tag{2}$$

$$E[Y_i] = \sum_{y_i=0}^{n} y_i \cdot \frac{n!}{y_i!(n - y_i)!} p_i^{y_i} (1 - p_i)^{n-y_i} \tag{3}$$

Expanding and simplifying this sum gives $E[Y_i] = np_i$.

Therefore, the expected value $E[Y] = \sum_{i=1}^{k} E[Y_i] = \sum_{i=1}^{k} np_i = n(p_1 + p_2 + p_3 + ... + p_k) = n$.

1.2 Given the following data

$$\left[\begin{array}{c|c} Age & Proportion \\ 18 - 24 & 0.18 \\ 25 - 34 & 0.23 \\ 35 - 44 & 0.16 \\ 45 - 64 & 0.27 \\ 65 - 100 & 0.16 \end{array}\right]$$

If 500 adults are sampled randomly, find the probability rhat the sample contains 100 person between 18 and 24, 200 between, 200 between the ages of 25 and 34, and 200 between the ages of 45 and 64. What is the expected value for to obtain a person in the 65 and above?

We can calculate the probability using the given proportions for each age group.

Let $p_{18-24} = 0.18$, $p_{25-34} = 0.23$, $p_{35-44} = 0.16$, $p_{45-64} = 0.27$, and $p_{65-100} = 0.16$.

We want to find the probability of obtaining 100 people between 18 and 24, 200 people between 25 and 34, 200 people between 45 and 64, and the remaining people aged 65 and above. The probability for this specific outcome is calculated using the multinomial probability formula

$$P = \frac{500!}{100! \cdot 200! \cdot 200! \cdot (500 - 100 - 200 - 200)!} \times (0.18)^{100} \times (0.23)^{200} \times (0.27)^{200} \times (0.16)^{500-100-200-200} \tag{4}$$

You can compute this expression to find the probability.

To find the expected value for obtaining a person aged 65 and above, you would use the proportion for that age group, which is $p_{65-100} = 0.16$, and multiply it by the total number of samples (500)

Expected value for 65 and above $= 0.16 \times 500 = 80$.

2. Consider the following experiment

   (a) The experiment consists of a fixed number, n, of identical trials.

   (b) Each trial results in one of two outcomes: success, S, or failure, F.

   (c) The probability of success on a single trial is equal to some value p and remains the same from trial to trial. The probability of a failure is equal to $q = (1 - p)$.

   (d) The trials are independent.

   (e) The random variable of interest is Y, the number of successes observed during the n trials.

2.1 From the following steps shown above, derive the probability distribution of the experiment.

The probability distribution of the experiment follows a binomial distribution, which can be represented as

$$P(Y_i = y_i) = \binom{n}{y_i} p^{y_i} (1-p)^{n-y_i} \tag{1}$$

where

- n is the number of trials,
- p is the probability of success on each trial,
- y_i is the number of successes observed during the trials, and
- $(1-p)$ is the probability of failure on each trial.

2.2 Show that the expectation of this probability distribution is $E[y] = \Sigma y \cdot P(y) = np$.

The expectation of a binomial distribution is given by

$$E[Y_i] = \sum_{y_i=0}^{n} y_i \cdot P(Y_i = y_i) = np \tag{2}$$

2.3 Experience has shown that 30% of all persons afflicted by a certain illness recover. A drug company has developed a new medication. Ten people with the illness were selected at random and received the medication; nine recovered shortly thereafter. Suppose that the medication was absolutely worthless. What is the probability that at least nine of ten receiving the medication will recover?

Given that the medication is absolutely worthless, the probability of success p is actually the same as the probability of recovery without the medication, which is 30% or 0.3. So, $p = 0.3$ and $q = 1 - p = 0.7$.
We want to find $P(Y \geq 9)$ when $n = 10$. We can calculate this using the binomial probability formula

$$\begin{aligned} P(Y \geq 9) &= P(Y = 9) + P(Y = 10) \\ &= \binom{10}{9} \times 0.3^9 \times 0.7^{10-9} + \binom{10}{10} \times 0.3^{10} \times 0.7^{10-10} \\ &= 10 \times 0.3^9 \times 0.7 + 1 \times 0.3^{10} \times 0.7^0 \\ &= 10 \times 0.3^9 \times 0.7 + 0.3^{10} \end{aligned} \tag{3}$$

So, the probability that at least nine out of ten individuals receiving the medication will recover is approximately 0.000144.

2.4 Suppose that a lot of 5000 electrical fuses contains 5% defectives. If a sample of 5 fuses is tested, find the probability of observing at least one defective.

Given that the lot contains 5% defectives, $p = 0.05$ and $q = 1 - p = 0.95$. We are selecting a sample of $n = 5$ fuses.

We want to find $P(Y \geq 1)$. Using the complement rule, we find $P(Y \geq 1) = 1 - P(Y = 0)$.

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1 - 0.05)^5. \tag{4}$$

So, the probability of observing at least one defective fuse in a sample of 5 is approximately 0.2262 or 22.62%.

3. Consider a probability distribution of $p(y) = \frac{\lambda^y}{y!} e^{-\lambda}$.

3.1 Find a general formula for the expected value of this distribution.

The expected value (or mean) of a probability distribution is given by the sum of each possible outcome multiplied by its probability. For this distribution, the possible outcomes are non-negative integers $(y = 0, 1, 2, ...)$. So, the expected value $E(y)$ is calculated as

$$E(y) = \sum_{y=0}^{\infty} y \cdot p(y) \tag{1}$$

where $p(y) = \frac{\lambda^y}{y!} e^{-\lambda}$. Let's evaluate this

$$\begin{aligned} E(y) &= \sum_{y=0}^{\infty} y \cdot \frac{\lambda^y}{y!} e^{-\lambda} \\ &= \sum_{y=1}^{\infty} y \cdot \frac{\lambda^y}{y!} e^{-\lambda} \quad \text{since } y = 0 \text{ yields 0} \\ &= e^{-\lambda} \sum_{y=1}^{\infty} \frac{\lambda^y}{(y-1)!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{k!} \quad \text{where } k = y - 1 \\ &= e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda e^{\lambda} = \lambda \quad \text{using } \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda} \\ &= \lambda \end{aligned} \tag{2}$$

So, the expected value of this distribution is simply λ.

3.2 A certain type of tree has seedlings randomly dispersed in a large area, with the mean density of seedlings being approximately five per square yard. If a forester randomly locates ten 1-square-yard sampling regions in the area, find the probability that none of the regions will contain seedlings.

Given that the mean density of seedlings is 5 per square yard, the parameter λ in our distribution is 5. We want to find the probability that none of the ten 1-square-yard sampling regions contain seedlings. This is the same as finding the probability that the number of seedlings in each region is 0.

Using the given probability distribution, with $\lambda = 5$. the probability of getting 0 seedlings in one region is

$$p(y = 0) = \frac{5^0}{0!} e^{-5} = e^{-5} \tag{3}$$

So, the probability that none of the regions contain seedlings is

$$P(\text{no seedlings in any region}) = (e^{-5})^1 0 \tag{4}$$

Therefore, the probability that none of the regions will contain seedlings is e^{-50}