Statistics and inferencing

Name: Basil Khowaja

Id: BK08432

"Activity 01"

a) Prove that the mean and variance of a bernoulli random variable are p and $p(1-p)$.

Ans// a bernoulli random variable x takes the value 1 with Probability p and the value 0 with Probability $1-p$. Therefore the PMF of x is :-

$$P(x=1) = p \quad \text{and} \quad P(x=0) = 1-p$$

$\Rightarrow$ the mean of x is:

$$\mu = E(x) = \sum_{x} x \cdot P(x=x)$$

for a bernoulli random variable:

$$E(x) = 1 \cdot P(x=1) + 0 \cdot P(x=0)$$

$$E(x) = 1 \cdot P + 0 \cdot (1-P) = P$$

So, the mean $\boxed{\mu = P}$

$\Rightarrow$ the variance of x is given by

$$\text{var}(x) = E(x^2) - (E(x))^2$$

Since x can only take values 0 or 1:

$$x^2 = x$$

So, $E(x^2) = E(x) = P$

Therefore,

$$\text{var}(x) = P - P^2 = P(1-P)$$

hence Proved.

b) Prove that the mean and variance of binomial random variable are np and $np(1-p)$ respectively.

Ans//

1) Mean of a Binomial random variable.

A binomial RV represents the number of Successes in n independent Bernoulli trials, where each trial has a Success Probability p. The PMF of x is given by:

$$P(x=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2 \dots n$.

The mean of x can be found using the following relationship

$$\mu = E(x) = \sum_{k=0}^{n} k \cdot P(x=k)$$

alternatively, Since X is the Sum of n independent Bernoulli random variables

$X_1, X_2, X_3 \dots X_n$ each with mean p, we have

$$E(x) = E\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} E(x_i)$$

$$= \sum_{i=1}^{n} p = np$$

So, the mean $\mu = np$.

$\Rightarrow$ variance of Binomial random variable:

The variance of x is given by:

$$Var(x) = E(x^2) - (E(x))^2$$

Since x is the sum of n independent Bernoulli random variables $X_1, X_2, X_3, \dots X_n$ and each x_i has variance $p(1-p)$, the variance of x is:

$$var(x) = var\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} var(x_i)$$

$$= \sum_{i=1}^{n} p(1-p) = np(1-p).$$

hence Proved.

c) if x_i, $1 \le i \le n$, are identical and independent Bernoulli distributed random variables, then convince yourself that $z = X_1 + X_2 + \dots X_n$ is binomially distributed.

Ans// $\Rightarrow$ the Probability that exactly k out of the n independent Bernoulli trials are successes (i.e. $z = k$) is given by the binomial distribution :

$$P(z = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$\Rightarrow$ this formula is derived from the fact that there are $\binom{n}{k}$ ways to choose which k trials out of n are successes, each occurring with Probability p^k, while the remaining $n-k$ trials are failures with Probability $(1-p)^{n-k}$.

$\Rightarrow$ the random variable $Z = X_1 + X_2 + \cdots X_n$ where X_i are identical and independent Bernoulli random variables, follows a Binomial distribution with Parameters n and P. This is because Z represents the total number of Successes in n independent trials, which is the defining characterstic of a binomial random variable.

d) let X_i and X_j be two bernoulli distributed random variables, find the moment $E[X_i X_j]$ under two assumptions

Ans//

case A) X_i and X_j are indelendent

$\Rightarrow$ when both of them are independent, the expected value of the Product $X_i X_j$ is given by the Product of their individual expectations

1) expectation of x_i:

$$E[x_i] = p_i$$

2) expectation of x_j:

$$E[x_j] = p_j$$

3) expectation of $x_i x_j$ when independent:

$$E[x_i x_j] = p_i \cdot p_j$$

case B) x_i and x_j are not independent

when they are not independent, we need to consider the joint probability distribution

1) Joint Probability:

- let $P(x_i = 1, x_j = 1) = p_{ij}$ represent the probability that both $x_i = 1$ and $x_j = 1$.

2) expectation of $x_i x_j$ when not independ.

$$E[x_i x_j] = P(x_i = 1, x_j = 1) = p_{ij}$$

$\Rightarrow$ in this case it's not necessary that

P_{ij} is equal to $P_i P_j$ since the variables are not independent

Summary:

① Independent case:

$$E[x_i x_j] = P_i \cdot P_j$$

② dependent case:

$$E[x_i x_j] = P_{ij}$$

e) expand the expression $(x_1 + x_2 + \cdots x_n)^2$ and express the terms in a compact form.

Ans// for expanding this we will apply the binomial expansion.

$$(x_1 + x_2 + \cdots x_n)^2$$

$$\hookrightarrow (x_1 + x_2 + \cdots x_n)(x_1 + x_2 \cdots x_n)$$

$\Rightarrow$ multiplying out each term:

$$= x_1^2 + x_1x_2 + \cdots + x_1x_n + x_2x_1 + x_2^2 + \cdots + x_2x_n + \cdots + x_nx_1 + x_nx_2 + \cdots$$

$\Rightarrow$ Grouping the terms:

we note that the each term x_i^2 appears once, and each Product x_ix_j where $i \neq j$ appears twice (as x_ix_j and x_jx_i)

So we can rewrite the expanded expression as:

$$\sum_{i=1}^{n} x_i^2 + 2 \sum_{1 \leq i < j \leq n} x_ix_j \Rightarrow \text{Ans}//$$

compact form:

$\Rightarrow$ the first sum $\sum_{i=1}^{n} x_i^2$ represents the sum of the squares of the individual terms.

$\Rightarrow$ the Second Sum $2\sum_{1\leq i<j\leq n} x_i x_j$ represents the Sum of all distinct Product terms $x_i x_j$ where i and j are distinct.