Dhanraj Babasoheb Bhosale 1225506620

Homework - 4

Q.1 $X_i = \begin{cases} 1 & \text{if } H \\ 0 & \text{if } T \end{cases}$

N denotes flip until we get 100 H

∴ Given $y = x_1 + x_2 + \cdots + x_N$

Since N and x_i are dependent, we can clearly see, ∴ y is not a random sum of random variable.

For any $N = n$, we know we will have 100 H.

∴ $y = x_1 + x_2 \ldots + x_n = 100$ always.

∴ pmf of y =

$$P_Y(y) = \begin{cases} 1 & y = 100 \\ 0 & \text{otherwise.} \end{cases}$$

Q.2 Given:

$P[V] = \frac{3}{4}$ $P[D] = 1/4$

a) $E[K_{100}]$

K_n is number of video packets in a collection of n packets.

∴ it follows binomial distribution

$$\therefore E[K_{100}] = 100\,E[K]$$
$$= 100\left[\frac{3}{4} \times 1 + \frac{1}{4} \times 0\right]$$

$$E[K_{100}] = 75$$

b) $\sigma_{K_{100}} = \left[E[K_{100}^2] - (E[K_{100}])^2 \right]^{1/2}$

For a binomial distribution.

$\sigma = \sqrt{npq} = \sqrt{100 \times \frac{3}{4} \times \frac{1}{4}}$

$\sigma_{K_{100}} = 4.33$

c) Central limit theorem to estimate

$P[K_{100} \geq 10]$

$P[K_{100} \geq 100] = 1 - P[K_{100} \leq 100]$

$= 1 - \phi\left(\frac{18-75}{4.33}\right) = 1 - \phi(-13.1639)$

$= 1 - (1 - \phi(13.1639))$

$= \phi(13.1639)$

$\therefore P[K_{100} \geq 100] = 1$

d) $P[16 \leq K_{100} \leq 24]$

$Z = \frac{K_{100} - 75}{4.33}$

$P[16 \leq K_{100} \leq 24] = \phi\left(\frac{24+0.5-75}{4.33}\right) - \phi\left(\frac{16+0.5-75}{4.33}\right)$

$= \phi(-11.662) - \phi(-13.510)$

$= 0.$

$\therefore$ This was using De-Moivre-Laplace approximation.

Q. 3 $Y_n = X_{2n-1} + (-X_{2n})$

a) $E[Y_n] = E[X_{2n-1} - X_{2n}]$

$= E[X_{2n-1}] - E[X_{2n}]$

$= E[X] - E[X] = 0$

{ This is according to theorem that expected value of sum equals sum of expected value.

$E[W_n] = E[X_1] + E[X_2] + \dots E[X_n]$

where $W_n = X_1 + \dots + X_n$ }

$Var[Y_n]$

according to theorem, variance of sum is sum of variances for uncorrelated $X_1, \dots X_n$.

$\therefore Var[Y_n] = Var[X_{2n-1}] + Var[-X_{2n}]$

$= 2\, Var[X]$

b) $M_n(Y)$

$E[M_n(Y)] = E[Y]$

according to sample mean theorem.

$\therefore E[M_n(Y)] = E[Y_n] = 0$

$Var[M_n(Y)] = \frac{Var[Y_n]}{n} = \frac{2\, Var[X]}{n}$

Q 4 Chebyshev's inequality states

$$P[|X - E[X]| \geq c] \leq \frac{\sigma_x^2}{c^2}$$

we substitute $c = k\sigma$

$$\Rightarrow P[|X - E[X]| \geq k\sigma] \leq \frac{1}{k^2}$$

Actual probability that Gaussian variable Y is more than k standard deviation from expected value is

$$P[|Y - E[Y]| \geq k\sigma_y] = P[Y - E[Y] \leq k\sigma_y] + P[Y - E[Y] \geq k\sigma_y]$$

$$= 2P\left[\frac{Y - E[Y]}{\sigma_y} \geq k\right]$$

$$= 2Q(k)$$

Comparison of result of upper bound

	$K=1$	$K=2$	$K=3$	$K=4$	$K=5$
Chebyshev's bound	1	0.250	0.111	0.0625	0.04
$2\Phi(k)$	0.317	0.046	0.0027	6.3×10^{-5}	5.7×10^{-7}

Chebyshev bound gets weak for higher 'k' values.

Q.5 To show, $M_n(X) = \frac{X_1 + \dots + X_n}{n}$

satisfies $\Rightarrow$

$$P[M_n(X) \geq c] \leq \left(\min_{s \geq 0} e^{-sc} \phi_x(s)\right)^n$$

Let $W_n = X_2 + \dots + X_n$

$$\therefore M_n(X) = \frac{W_n}{n}$$

$$\therefore P[M_n(X) \geq c] = P[W_n \geq nc]$$

$$\phi_{W_n}(s) = (\phi_x(s))^n$$

applying chernoff bound to $W_n \Rightarrow$

$$P[W_n \geq nc] \leq \min_{s \geq 0} e^{-snc} \phi_{W_n}(s)$$

$$P[W_n \geq nc] = \min_{s \geq 0} (e^{-sc} \phi_x(s))^n$$

for $y \geq 0$, y^n is increasing

$\therefore$ Value of s that minimizes $e^{-sc} \phi_x(s)$ also minimize $(e^{-sc} \phi_x(s))$

$$P[M_n(X) \geq c] = P[W_n \geq nc] \leq \left(\min_{s \geq 0} e^{-sc} \phi_x(s)\right)^n$$

Q 6 $P[A] = 0.8$

Since X_A is bernoulli

$$E[X_A] = 1 \cdot P[A] + 0 \cdot (1 - P[A])$$

$$E[X_A] = 0.8$$

$$Var[X_A] = P[A](1 - P[A]) = 0.16$$

b) $Var[\hat{P}_n(A)]$

Let $X_{A,i}$ denote X_A for ith trial

we know,

$$\hat{P}_n(A) = M_n(X_A) = \frac{1}{n} \sum_{i=1}^{n} X_{A,i}$$

$$\therefore Var[\hat{P}_n(A)] = \frac{1}{n^2} \sum_{i=1}^{n} Var[X_{A,i}]$$

$$= \frac{P[A](1 - P[A])}{n}$$

$$= \frac{0.16}{n}$$ after substitution

c) $P[|\hat{P}_n(A) - P[A]| \leq 0.1] \geq 1 - \alpha$

$\therefore \hat{P}_{100}(A) = M_{100}(X_A)$

we can write

$$P[|\hat{P}_{100}(A) - P[A]| < c] \geq 1 - \frac{Var(X_A)}{100 c^2}$$

$$= 1 - \frac{0.16}{100 c^2} = 1 - \alpha$$

for $c = 0.1$

$$\alpha = \frac{0.16}{100(0.1)^2} = 0.16$$

Thus for 100 samples, confidence coefficient

$$\Rightarrow 1 - \alpha = 0.84$$

d) find n, such that

$$P[|\hat{P}_n(A) - P[A]| \leq 0.1] \geq 0.95$$

$$P[|\hat{P}_n(A) - P[A]| < c] \geq 1 - \frac{var[X_n]}{n c^2}$$

$$1 - \frac{0.16}{nc^2} = 1 - \alpha$$

$\therefore$ for $c = 0.1$ we have confidence coefficient $1 - \alpha = 0.95$

$$\text{if } \alpha = \frac{0.16}{n(0.1)^2} = 0.05$$

or $\underline{n = 320}$