Justin Li
4/27/23

COEN 239 HW 3

1. The response time of a computer system has an Erlang distribution with the following cumulative distribution function (CDF):

$$F(x) = 1 - e^{-x/a}\{(\text{from } i=0 \text{ to } m-1)\ (x=a)^i/i!\ \},\ x >= 0$$

Find expressions for the pdf, mean, variance, mode, and coefficient of variation of the response time.

pdf:

$$f(x) = \frac{dF(x)}{dx} = -e^{-x/a}\left(-\frac{1}{a}\right)\sum_{i=0}^{m-1}\frac{(x/a)^i}{i!} - e^{-x/a}\sum_{i=1}^{m-1}\frac{i x^{i-1}}{i!\, a^i}$$

$$= \frac{e^{-x/a}}{a}\left\{\sum_{i=0}^{m-1}\frac{(x/a)^i}{i!} - \sum_{i=1}^{m-1}\frac{x^{i-1}}{a^{i-1}}\cdot\frac{1}{(i-1)!}\right\}$$

$$= \frac{e^{-x/a}}{a}\left\{1 + \frac{(x/a)}{1!} + \frac{(x/a)^2}{2!} + \cdots + \frac{(x/a)^{m-2}}{(m-2)!} + \frac{(x/a)^{m-1}}{(m-1)!} - 1 - \frac{(x/a)}{1!} - \frac{(x/a)^2}{2!} - \cdots - \frac{(x/a)^{m-2}}{(m-2)!}\right\}$$

$= x^{m-1} e^{-x/a}/(m-1)!a^m,\ x >= 0$

Mean:

$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^{\infty}\frac{x^m e^{-x/a}}{a^m (m-1)!}\,dx$$

$$= am\underbrace{\int_0^{\infty}\frac{x^m e^{-x/a}}{a^{m+1}\, m!}\,dx}_{=1} = am$$

Variance:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^{\infty} \frac{x^{m+1} e^{-x/a}}{a^m (m-1)!}\,dx$$

$$= a^2 (m+1) m \underbrace{\int_0^{\infty} \frac{x^{m+1} e^{-x/a}}{a^{m+2} (m+1)!}\,dx}_{=1} = a^2 m (m+1)$$

$$VAR(X) = E(X^2) - \mu^2 = a^2 m (m+1) - (am)^2$$

$$= a^2 m (m+1-m) = a^2 m$$

Mode:

$$\frac{df(x)}{dx} = \frac{e^{-x/a}}{a^m (m-1)!} \left[-\frac{1}{a} x^{m-1} + (m-1) x^{m-2} \right]$$

$$= \frac{e^{-x/a} x^{m-2}}{a^m (m-1)!} \left[(m-1) - \frac{x}{a} \right] = 0$$

$\frac{df(x)}{dx} = 0$ OCCURS AT $x = 0$, AND $x = a(m-1)$

Coefficient of Variation:

$$C_v = \frac{\sqrt{a^2 m}}{am} = \frac{1}{\sqrt{m}}$$

2. The cumulative distribution function (CDF) of a Pareto variate is given by

   $F(x) = 1 - x^{-a}$; $1 < x < \inf$

   Find expressions for pdf, mean, variance, mode, and coefficient of variation.

PDF : $f(x) = \frac{dF(x)}{dx} = \frac{a}{x^{a+1}}$; $x \geq 1$

$$f(x) = \begin{cases} a x^{-a-1}, & x \geq 1 \\ 0 , & \text{ELSE WHERE} \end{cases}$$

$E(X)$ = MEAN OF RV $X \triangleq \mu$

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = a \int_{1}^{\infty} \frac{dx}{x^a}$$

$a = 1$; $E(X) = a \ln x \Big|_1^{\infty} \rightarrow \infty \Rightarrow$

$E(X)$ DOES NOT EXIST

$a \neq 1$; $E(X) = \frac{a x^{-a+1}}{-a+1} \Big|_1^{\infty} = -\frac{a}{(a-1)} \frac{1}{x^{a-1}} \Big|_1^{\infty}$

$= \frac{a}{a-1}$ IF $a > 1$

$= \infty$ IF $a < 1$ $\Rightarrow$ DOES NOT EXIST

$$E(X) = \begin{cases} \text{DOES NOT EXIST} & a \leq 1 \\ \frac{a}{a-1} & a > 1 \end{cases}$$

Variance:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \text{SECOND MOMENT OF RV } X$$

$$E(X^2) = \int_{1}^{\infty} x^2 \cdot \frac{a}{x^{a+1}}\,dx = a\int_{1}^{\infty} x^{-a+1}\,dx$$

$$= a\,\frac{x^{-a+2}}{-a+2}\Bigg|_{1}^{\infty} = \frac{a}{a-2} \quad \text{IF } a > 2$$

$$\text{VAR}(X) = E(X^2) - \mu^2 \triangleq \sigma_X^2$$

$$\sigma_X^2 = \frac{a}{(a-2)} - \left(\frac{a}{a-1}\right)^2 = \frac{a}{(a-1)^2(a-2)} \qquad a > 2$$

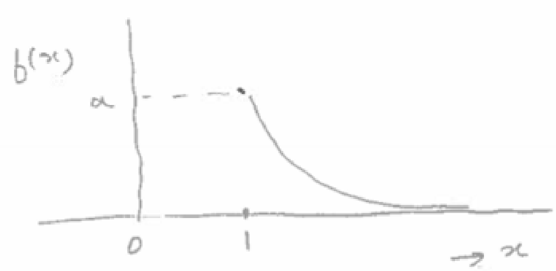
$$\text{COEFFICIENT OF VARIATION} = C_v = \frac{\sigma_X}{|\mu|}$$

$$= \frac{\sqrt{a}}{(a-1)(a-2)^{1/2}}\,\frac{(a-1)}{a}$$

$$\therefore C_v = \frac{1}{\sqrt{a(a-2)}} \;; \quad a > 2$$

Mode:

PLOT OF $f(x)$



$f(x)$ IS MAXIMUM AT $x = 1$

$\therefore$ MODE OCCURS AT $x = 1$

3. The execution times of queries on a database is normally distributed with a mean of 5 seconds and a standard deviation of one second. Determine the following:
    a. What is the probability of the execution time being more than 8 seconds?
    0.0013
    b. What is the probability of the execution time being less than 6 seconds?
    0.8413
    c. What percentage of responses will take between 4 and 7 seconds?
    0.8185
    d. What is the 95-percentile execution time?
    6.664 sec

4. What index of central tendency should be used to report?
    a. Response time (symmetrical pdf)
    Mean
    b. Number of packets per day (symmetrical pdf)
    Mean
    c. Number of packets per second (skewed pdf)
    Median
    d. Frequency of keywords
    Mode

5. How would you summarize an average personal computer configuration:
    a. CPU type
    Mode
    b. Memory size
    Median
    c. Disk type
    Mode
    d. Number of peripherals
    Median
    e. Cost
    Median

6. The CPU times in milliseconds for 11 workloads on a processor are: 0.74, 0.43, 0.24, 2.24, 262.08, 8960, 4720, 19740, 7360, 22440, and 28560. Which index of central tendency would you choose and why?
    Since the ratio of maximum to minimum is very high, use the median.

7. The number of disk I/Oís performed by a number of programs were measured as follows: {23,33, 14, 15, 42, 28, 33, 45, 23, 34, 39, 21, 36, 23, 34, 36, 25, 9, 11, 19, 35, 24, 31, 29, 16, 23, 34, 24, 38, 15, 13, 35, 28}. Which index of central tendency would you choose and why?
    Mean since the data is very clustered together (not skewed) and y_{max} = y_{min} ratio is small.

Problem 2: You are given an undirected graph $G = (V; E)$, where elements of the vertex set V are 1, 2, 3, and 4: The edge set E is given by $\{v, w, x, z\}$ where $v = (1, 3)$; $w = (2, 3)$; $x = (2, 4)$; and $z = (1, 4)$.

- Using the matrix-tree theorem, prove that the number of spanning trees is 4.

$$A = \begin{matrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{matrix} \qquad D = \begin{matrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{matrix} \qquad M = \begin{matrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{matrix}$$

$$(4,3) \text{ Cofactor of } M = (-1)^{4+3} \begin{matrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 0 \end{matrix} = -\{2(-1) - 2\} = -\{-4\} = 4$$

- Using the generalized form of matrix tree theorem, list all the spanning trees.

$$A' = \begin{matrix} 0 & 0 & v & z \\ 0 & 0 & w & x \\ v & w & 0 & 0 \\ z & x & 0 & 0 \end{matrix} \qquad D' = \begin{matrix} v+z & & & \\ & w+x & & \\ & & v+w & \\ & & & z+x \end{matrix} \qquad M' = \begin{matrix} v+z & 0 & -v & -z \\ 0 & w+x & -w & -x \\ -v & -w & v+w & 0 \\ -z & -x & 0 & z+x \end{matrix}$$

$$(4,3) \text{ Cofactor of } M' = (-1)^{4+3} \begin{matrix} v+z & 0 & -z \\ 0 & w+x & -x \\ -v & -w & 0 \end{matrix}$$

$= -\{(v+z)(-wx) - z(vw+vx)\}$

$= vwx + wxz + vwz + vxz$

Problem 3: Consider a three node fully connected undirected graph. The degree of each node in this graph is two. The probability that each link is operational is p.

- List all the 3 distinct spanning trees of this graph.
  $t1 = \{a, b\}$, $t2 = \{b, c\}$, $t3 = \{a, c\}$
- Using the principle of inclusion and exclusion theorem of probability theory, prove that the reliability of this network is given by $p^2 (1 + 2q)$; where $q = (1 - p)$.
  $P(T1) = P(T2) = P(T3) = p^2$
  $P(T1 \cap T2) = P(T2 \cap T3) = (T1 \cap T3) = p^3$
  $P(T1 \cap T2 \cap T3) = p^3$
  Using the inclusion and exclusion principle we obtain R(G) to be
  $R(G) = P(T1 \cup T2 \cup T3)$
  $= P(T1) + P(T2) + P(T3) - P(T1 \cap T2) - (T1 \cap T3) - P(T2 \cap T3) + P(T1 \cap T2 \cap T3)$
  $= 3p^2 - 3p^3 + p^3$
  $= 3p^2 - 2p^3$
  $= p^2(3 - 2p)$
  Substituting $p = (1 - q)$ in $(3 - 2p)$, we get

$$R(G) = p^2(1 + 2q)$$