

ON INEQUALITY TO GENERATE SOME STATISTICAL DISTRIBUTIONS

A. F. OGUNTOLU*¹, U. Y. ABUBAKAR¹, A. ISAH¹, L. A. NAFIU¹ AND K. RAUF²

ABSTRACT. In this work, we established Markov inequality via Binomial, Poisson and Geometric Distribution. Results obtained were used to obtain probability bound for some random variables. Our results are in agreement with the existing works.

1. INTRODUCTION

Markov inequality states that:

For any nonnegative random variable X and $a > 0$, we have the following inequality

$$\Pr(X \geq a) \leq \frac{E(X)}{a}.$$

This is equivalent to the following inequality in measure theory:

$$\mu(\{x \in X : |f(x)| \geq \epsilon\}) \leq \frac{1}{\epsilon} \int_X |f| d\mu.$$

Where (X, Σ, μ) is a measure space, f is a measurable extended real-valued function, and $\epsilon \geq 0$.

This definition is sometimes referred to as Chebyshev's or Bienayme's inequality in analysis because it appeared without prove in, at least, the work of Pafnuty Chebyshev who happened to be Markov's teacher (Stein and Shakarchi).

Markov's inequality (Markov (1884)) is useful in providing an upper bound for the probability of a random variable with a non negative function which is greater than or equal to some positive constant. It also relate probabilities to expectations.

In particular, it is applicable in proving Chebyshev's inequality (Chebyshev (1874)) and in showing that for a nonnegative random variable, the mean μ and a median m are such that $m \leq 2\mu$.

In this work, we provide a simpler and better proof to Markov Inequality through Binomial, Poisson and Geometric Distributions. Some applications are also given. For further work

*Corresponding author

2010 Mathematics Subject Classification: 41A50

Key words and phrases: Markov Inequality, Probability Distributions, Continuous random variable, Bound, Measure Space and Sharp

see Salo-Coste (1997), Ross (1976), Steliga and Szyal (2010), Feller (1966), Papoulis (1991), Clarkson et.al. (2009), DasGupta (2000), McWilliams (1990), Pitman (1993), Pfeiffer and Schum (1973).

2.0 Materials and Methods

This section consider the prove of Markov inequality and its application to some properties of some distributions.

2.1 The prove of Markov Inequality

THEOREM: Let X be a random variable and assume that its range is a subset of non-negative real numbers. Assume that $E[X]$ exists. Furthermore let A be some positive constant, then,

$$\Pr(X \geq A) \leq \frac{E[X]}{A}$$

Proof (Discrete): Let X be a discrete random variable. The proof is almost trivial; we only have to recall the notion of the mean of a random variable. It is

$$E[X] = \sum_i Z_i \Pr(X = Z_i)$$

Where Z_i is the sequence of the range of X . The right hand side contains just non-negative elements, thus we decrease this sum if we restrict it to those $Z_i \geq A$

$$E[X] = \sum_i Z_i * \Pr(X = Z_i) \geq \sum_{z_i \geq A} Z_i * \Pr(X = Z_i) \geq \sum_{z_i \geq A} A * \Pr(X = z_i) \geq A * \sum_{z_i \geq A} \Pr(X = z_i)$$

We can rewrite

$$\sum_{z_i \geq A} \Pr(X = z_i) \text{ as } \Pr(X \geq A) ,$$

$E[X] \geq A * \Pr(X \geq A)$ and

$$\Pr(X \geq A) \leq \frac{E[X]}{A}$$

Proof (Continuous): Let x be a continuous random variable, then,

$$\Pr(X \geq t) = \int_t^{\infty} \ell^{-x} dx$$

Let $U = x$ then $du = dx$ and $v = \ell^{-x}$ then $dv = -\ell^{-x}$. Hence,

$$\Pr(X \geq t) = x \ell^{-x} \Big|_t^{\infty} + \int_t^{\infty} x \ell^{-x} dx$$

$$= x \ell^{-x} \Big|_t^{\infty} + E[X \geq t]$$

$$= t \ell^{-x} + E[X \geq t]$$

$$= t \ell^{-x} + E[X] - E[X < t]$$

$$\Rightarrow \Pr(X \geq t) \leq \frac{E[X]}{t}$$

2.2 Properties of Binomial Distribution

The Binomial Distribution is given by

$$\Pr[X] = \binom{n}{x} P^x (1 - P)^{n-x}$$

While the Cumulative Probability is

$$\begin{aligned} \Pr[X \geq 0] &= \sum_{x=0}^n \binom{n}{x} P^x (1 - P)^{n-x} \\ &= \sum_{x=0}^n \frac{n!}{(n-x)! x!} P^x (1 - P)^{n-x} \\ &= n! \left[\frac{(1 - P)^n}{n! 0!} + \frac{P(1 - P)^{n-1}}{(n-1)!(1)!} + \dots + \frac{P^n}{0! n!} \right] \end{aligned}$$

By Binomial expansion, we have

$$\Pr[X \geq 0] = \frac{n!}{n!} \left[(1 - P)^n + nP(1 - P)^{n-1} + \dots + nP^{n-1}(1 - P) + P^n \right]$$

$$\Pr[X \geq 0] = (P + 1 - P)^n = 1^n = 1$$

$$E[X] = \sum_{x=0}^n \frac{x! n!}{(n-x)! x(x-1)!} P^x (1 - P)^{n-x}$$

$$\begin{aligned}
&= \sum_{x=0}^n \frac{n!}{(n-x)!(x-1)!} P^x (1-P)^{n-x} \\
&= \sum_{x=0}^n \frac{n(n-1)!}{(n-x)!(x-1)!} P \cdot P^{x-1} (1-P)^{n-x} \\
&= nP \sum_{x=0}^n \frac{(n-1)!}{(n-x)!(x-1)!} P^{x-1} (1-P)^{n-x}
\end{aligned}$$

$$\text{Mean} = E[X] = nP(1) = nP$$

$$\begin{aligned}
E[X^2] &= \sum_{x=0}^n \frac{x^2 n!}{(n-x)! x!} P^x (1-P)^{n-x} \\
&= \sum_{x=0}^n \frac{x n!}{(n-x)!(x-1)!} P^x (1-P)^{n-x} \\
&= \sum_{x=0}^n \frac{(x-1)n!}{(n-1)!(x-1)(x-2)!} P^x (1-P)^{n-x} + \sum_{x=0}^n \frac{n!}{(n-x)!(x-1)!} P^x (1-P)^{n-x} \\
&= \sum_{x=0}^n \frac{n(n-1)(n-2)!}{(n-x)!(x-2)!} P \cdot P \cdot P^{x-2} (1-P)^{n-x} + \sum_{x=0}^n \frac{n(n-1)!}{(n-x)!(x-1)!} P \cdot P^{x-1} (1-P)^{n-x} \\
&= n(n-1)P^2 \sum_{x=0}^n \frac{(x-2)!}{(n-x)!(x-2)!} P^{x-2} (1-P)^{n-x} + nP \sum_{x=0}^n \frac{(n-x)!}{(n-x)!(x-1)!} P^{x-1} (1-P)^{n-x} \\
&= n(n-1)P^2(1) + nP(1) \\
\Rightarrow E[X^2] &= n(n-1)P^2 + nP
\end{aligned}$$

$$\text{Variance } (\sigma^2) = \text{Var}[x] = E[X^2] - (E[X])^2$$

$$\begin{aligned}
\sigma^2 &= n(n-1)P^2 + nP - (nP)^2 \\
&= n^2 P^2 - nP^2 + nP - n^2 P^2
\end{aligned}$$

We have the variance as:

$$\text{Variance } (\sigma^2) = nP(1-P)$$

2.3 Prove of Markov Inequality from Binomial distribution

$$\begin{aligned}
\text{Pr}[X \geq a] &= \sum_{x=0}^n \binom{n}{x} P^x (1-P)^{n-x} \\
&= \frac{x}{x} \sum_{x=0}^n \binom{n}{x} P^x (1-P)^{n-x} \\
&\leq \frac{1}{x} \sum_{x=0}^n x \cdot \binom{n}{x} P^x (1-P)^{n-x} = \frac{1}{x} (nP)
\end{aligned}$$

But, $x \geq a \Rightarrow \frac{1}{x} \leq \frac{1}{a}$. Hence,

$$\Pr[X \geq a] \leq \frac{1}{a}(nP) = \frac{nP}{a}$$

2.4 Poisson Distribution

The Poisson distribution is given by

$$\Pr[X] = \frac{\lambda^x \ell^{-\lambda}}{x!}$$

Cummulative Probability

$$\Pr[X \geq 0] = \sum_{x=0}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{x!} = \ell^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \ell^{-\lambda} \ell^{\lambda} = 1$$

$$\begin{aligned} E[X] &= \sum_{x=0}^{\infty} \frac{x \cdot \lambda^x \ell^{-\lambda}}{x!} = \sum_{x=0}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{(x-1)!} \\ &= \lambda \sum_{x=0}^{\infty} \frac{\lambda^{x-1} \ell^{-\lambda}}{(x-1)!} = \lambda \ell^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda \ell^{-\lambda} \cdot \ell^{\lambda} = \lambda \end{aligned}$$

$$\text{Mean} = E[X] = \lambda$$

$$\begin{aligned} E[X^2] &= \sum_{x=0}^{\infty} \frac{x^2 \cdot \lambda^x \ell^{-\lambda}}{x!} = \sum_{x=0}^{\infty} \frac{x \cdot \lambda^x \ell^{-\lambda}}{(x-1)!} \\ &= \sum_{x=0}^{\infty} \frac{(x-1) \cdot \lambda^x \ell^{-\lambda}}{(x-1)!} + \sum_{x=0}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{(x-1)!} \\ &= \sum_{x=0}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{(x-2)!} + \lambda \sum_{x=0}^{\infty} \frac{\lambda^{x-1} \ell^{-\lambda}}{(x-1)!} \\ &= \lambda^2 \ell^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda \ell^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \end{aligned}$$

$$E[X^2] = \lambda^2 \ell^{-\lambda} + \lambda \ell^{-\lambda} \cdot \ell^{\lambda} = \lambda^2 + \lambda$$

$$\text{Variance } (\sigma^2) = \text{Var} [X] = E[X^2] - (E[X])^2$$

$$= \text{Var}[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Hence, we have the variance as:

$$\sigma^2 = \lambda$$

2.5 The prove of Markov's Inequality from Poisson distribution

$$\begin{aligned} \Pr[X \geq a] &= \sum_{x=a}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{x!} = \frac{\lambda}{x} \sum_{x=a}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{x!} \\ &= \frac{1}{x} \sum_{x=a}^{\infty} \frac{x \cdot \lambda^x \ell^{-\lambda}}{x!} = \frac{1}{x} \sum_{x=a}^{\infty} \frac{\lambda^x \ell^{-\lambda}}{(x-1)!} \\ &= \frac{\lambda}{x} \sum_{x=a}^{\infty} \frac{\lambda^{x-1} \ell^{-\lambda}}{(x-1)!} \leq \frac{\lambda}{x} \sum_{x=0}^{\infty} \frac{\lambda^{x-1} \ell^{-\lambda}}{(x-1)!} = \frac{\lambda}{x} \end{aligned}$$

$$\text{If } X \geq a \Rightarrow \frac{1}{X} \leq \frac{1}{a}$$

$$\text{Hence, } \Pr[X \geq a] \leq \frac{\lambda}{a}$$

2.6 Geometric Distribution

The Geometric Distribution is given by

$$\Pr[X] = P(1 - P)^x$$

The Cumulative Probability is

$$\sum_{x=0}^{\infty} P(1 - P)^x = P \sum_{x=0}^{\infty} (1 - P)^x$$

$$\text{Let } a = 1 - P \Rightarrow a - 1 = -P$$

$$\sum_{x=0}^{\infty} P(1 - P)^x = P \sum_{x=0}^{\infty} a^x$$

$$\sum_{x=0}^{\infty} a^x = \frac{a^{x+1} - 1}{a - 1} = \frac{1 - a^{x+1}}{1 - a} = \frac{P(1 - a^{x+1})}{(1 - a)}$$

$$\sum_{x=0}^{\infty} \Pr[X] = P \sum_{x=0}^{\infty} a^x = \frac{P(a^{x+1} - 1)}{-P} = -(a^{x+1} - 1) = (1 - a^{x+1}) = [1 - (1 - P)^{x+1}]$$

As $x + 1 \rightarrow \infty$ where $0 < P < 1$, $(1 - P) \rightarrow 0$

$$\sum_{x=0}^{\infty} \Pr[X] = (1 - 0) = 1$$

Expected value is

ON INEQUALITY TO GENERATE SOME STATISTICAL DISTRIBUTIONS

$$\sum_{x=0}^{\infty} x \cdot \Pr[X] = \sum_{x=0}^{\infty} x \cdot P (1 - P)^x$$

Let $a = 1 - P$

$$\sum_{x=0}^{\infty} x \cdot \Pr[X] = P \sum_{x=0}^{\infty} x \cdot a^x$$

$$\sum_{x=0}^{\infty} x \cdot \Pr[X] = P \left[\frac{a}{1-a} (1 - a^x) + \frac{a^2 (1-a)}{1-a} + \dots + \frac{a^{x-1} (1-a^2)}{1-a} + a^x \right]$$

$$\sum_{x=0}^{\infty} x \cdot \Pr[X] = \frac{P}{1-a} [a(1 - a^x) + a^2(1 - a^{x-1}) + \dots + a^{x-1}(1 - a^2) + a^x]$$

Since $P = 1 - a$

$$E[X] = a[(1 - a^x) + a(1 - a^{x-1}) + \dots + a^{x-2}(1 - a^2) + a^{x-1}(1 - a)]$$

$$= a \left(\frac{1 - a^x}{1 - a} \right) - xa^x$$

As $x \rightarrow 0$ within $0 < P < 1$ for $a = 1 - P$

$$a^x = (1 - P)^x \rightarrow 0 \text{ since } 0 < (1 - P) < 1$$

$$E[X] = a \left(\frac{1 - 0}{1 - a} - 0 \right) = a \left(\frac{1}{1 - a} \right) = \frac{a}{1 - a}$$

$$= \frac{1 - P}{1 - (1 - P)} = \frac{1 - P}{P}$$

$$\text{Mean} = E[X] = \frac{1 - P}{P}$$

The Variance is,

$$E[X^2] = \sum_{x=0}^{\infty} P \cdot x^2 \cdot a^x \quad \text{where } a = 1 - P$$

$$E[X^2] = P \sum_{x=0}^{\infty} x^2 \cdot a^x$$

$$= P[(1^2 - 0^2)a \left(\frac{1 - a^x}{1 - a} \right) + (2^2 - 1^2)a^2 \left(\frac{1 - a^x}{1 - a} \right) + \dots + [(x-1)^2 - (x-2)^2]a^{x-1} \left(\frac{1 - a^2}{1 - a} \right) +$$

$$(x^2 - (x-1)^2)a^x]$$

$$= a[(1^2 - 0^2) + (2^2 - 1^2)a + \dots + [(x-1)^2 - (x-2)^2]a^{x-2} + x^2 - (x-1)^2]a^{x-1} -$$

$$[(1^2 - 0^2) + (2^2 - 1^2) + \dots + [(x-1)^2 - (x-2)^2] + [x^2 - (x-1)^2]]a^x]$$

$$= a \left[\left(2 \sum_{x=0}^{\infty} xa^x + \sum_{x=0}^{\infty} a^x \right) - \left[a^x \sum_{x=0}^{\infty} (2x-1) \right] \right]$$

$$\text{But } \sum_{x=0}^{\infty} xa^x = \frac{a}{1-a} \left[\frac{1-a^x}{1-a} - xa^x \right] \quad \text{And } \sum_{x=0}^{\infty} (2x-1)a^x = x^2$$

$$E[X^2] = a \left[\frac{2a}{1-a} \left(\frac{1-a^x}{1-a} - xa^x \right) + \left(\frac{1-a^{x+1}}{1+a} \right) \right] - x^2 \cdot a^{x+1}$$

As $x \rightarrow \infty$, $a^x \rightarrow 0$

$$\begin{aligned} E[X^2] &= a \left[\left(\frac{2a}{1-a \left(\frac{1}{1-a} - 0 \right)} + \frac{1}{1-a} \right) - 0 \right] \\ &= \frac{2a^2}{(1-a)^2} + \frac{a}{1-a} = \frac{2a^2 + a(1-a)}{(1-a)^2} = \frac{a^2 + a}{(1-a)^2} \\ &= \frac{(1-P)^2 + (1-P)}{P^2} \end{aligned}$$

We have the variance as:

$$\text{Variance} = \sigma^2 = \text{Var}[X] = E[X^2] - (E[X])^2$$

$$\sigma^2 = \frac{(1-P)^2 + (1-P)}{P^2} - \frac{(1-P)^2}{P^2} = \frac{(1-P)}{P^2}$$

2.7 Prove of Markov's Inequality from Geometric distribution

$$\begin{aligned} \Pr[X \geq a] &= \sum_{x=a}^{\infty} P(1-P)^x \\ &= \frac{x}{x} \sum_{x=a}^{\infty} P(1-P)^x \\ &\leq \frac{1}{x} \sum_{x=a}^{\infty} x \cdot P(1-P)^x = \frac{1}{x} E[X] = \frac{1}{x} \left(\frac{1-P}{P} \right) \end{aligned}$$

$$\text{But } X \geq a \Rightarrow \frac{1}{X} \leq \frac{1}{a}$$

$$\Pr[X \geq a] \leq \frac{E[X]}{a} \leq \frac{1-P}{aP}$$

3. Tightness of Markov's Inequality

Markov Inequality provides a weak bound because the amount of information provided to it is limited. All the information provided to it were that the variable is non-negative and that the expected value is known and finite. Now, we will show that it is indeed tight-that is Markov Inequality is already doing as much as it can. Consider a random variable X such that

$$X = \begin{cases} k & \text{with prob. } \frac{1}{k} \\ 0 & \text{else} \end{cases}$$

We can estimate its expected value as:

$$E[X] = \frac{1}{k} * k + \frac{k-1}{k} * 0 = 1$$

$$\Pr(X \geq kE[X]) = \Pr(X \geq k) = \frac{1}{k}$$

This implies that the bound is tight because the value of the random variable is exactly k or otherwise zero.

4. Results

We use Markov inequality to analyze Binomial distribution, Poisson distribution and Geometric distribution given that the initial probability distribution is $P = \frac{1}{a}$ where a is a positive natural number .

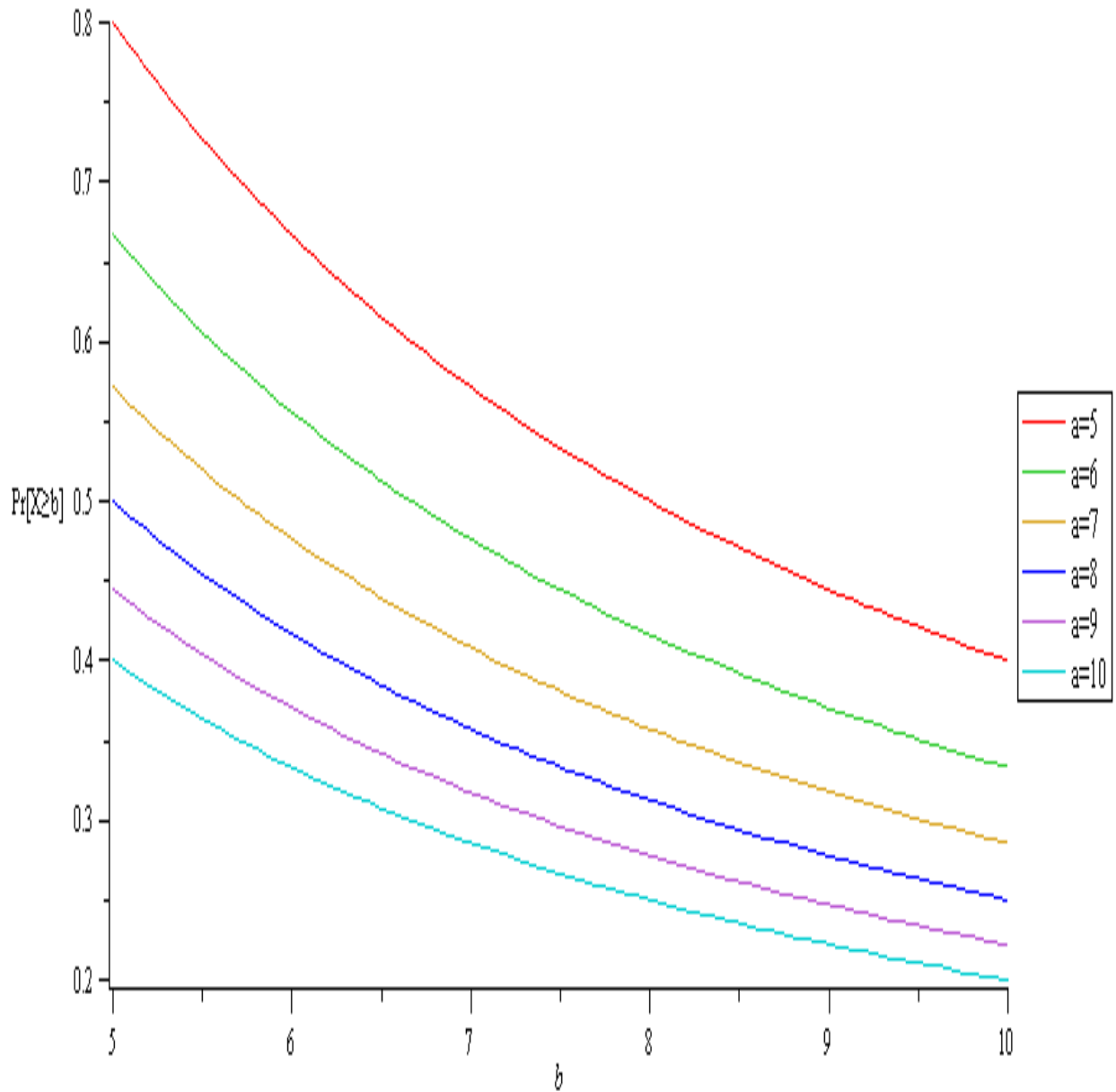
By plotting the graph of Markov Inequalities. We have the following:

Markov Inequality from binomial distribution

$$\Pr[X \geq b] \leq \frac{E[X]}{b} \leq \frac{np}{b}. \text{ But, } p = \frac{1}{a} \text{ and } \Pr[X \geq b] \leq \frac{n}{a.b}$$

We consider the equation for the graph as: $\Pr[X \geq b] \leq \frac{n}{a.b}$

Figure 1



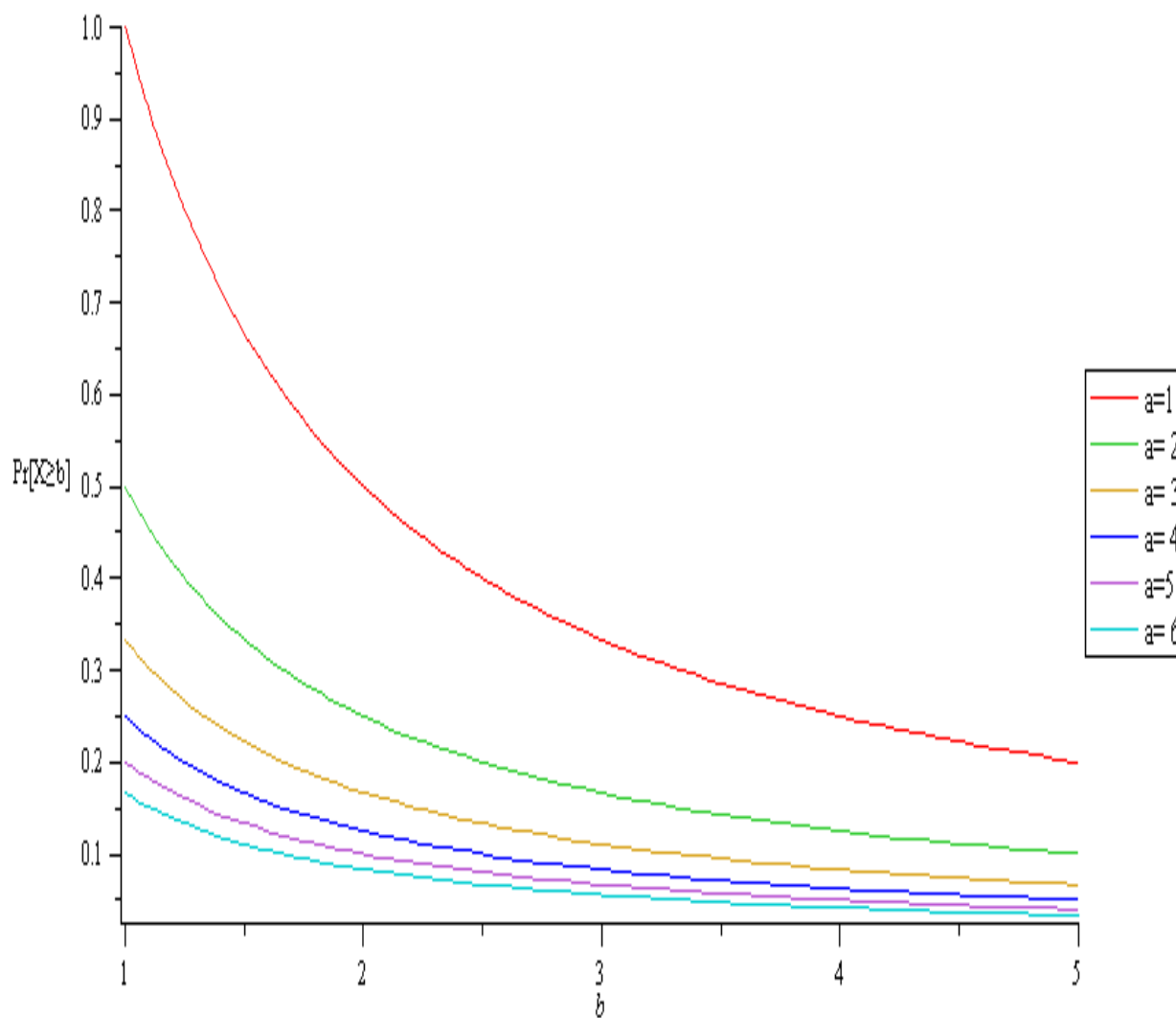
The graph of $\Pr[X \geq b]$ againsts number of trials b

Markov Inequality from Poisson distribution

$$\Pr[X \geq b] \leq \frac{\lambda}{b}$$

$$\text{But } \lambda = \frac{1}{a} \text{ and } \Pr[X \geq b] \leq \frac{1}{a \cdot b}$$

Figure 2

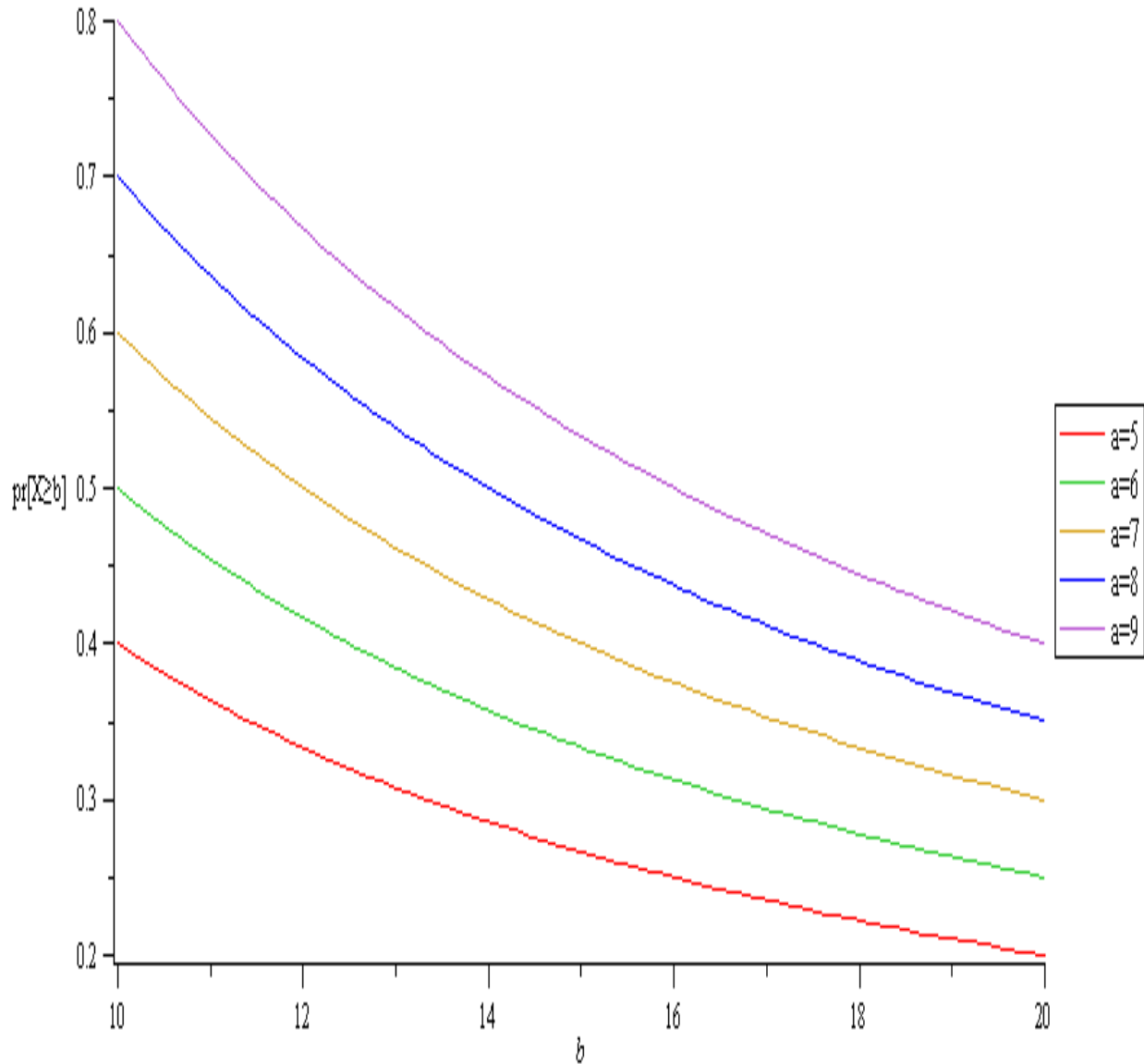


The graph of $\Pr[X \geq b]$ against number of trials b

Markov Inequality from Geometric distribution

$$\Pr[X \geq b] \leq \frac{(1-p)}{p \cdot b} = \frac{(1 - \frac{1}{a})}{\frac{1}{a} \cdot b} = \frac{a-1}{b}$$

Figure 3



The graph of $\Pr[X \geq b]$ against number of trials b

4 Discussions

In Figure 1, we observe that with increase in the value of ' a ' the probability steadily reduces. Also, it can be observed that as the number of trials increases, the probability of the event will considerably reduce.

In Figure 2, we observe that with increase in the value of ' a ', the probability steadily reduces. Also, it can be observed that as the number of trials increases, the probability of the event will considerably reduce.

In Figure 3, we observe that with increase in the value of ' a ', the probability steadily reduces. Also, it can be observed that as the number of trials increases, the probability of the event will considerably reduce.

5 Conclusions

The use of Markov inequality to analyze the Binomial distribution is inadequate because it gives unrealistic inequality variation, thus making it difficult to properly establish a realistic bound for the probability of higher value, hence a new approach is required to establish a more realistic bound.

The use of Markov inequality to analyze the Poisson distribution is adequate because it gives a suitable inequality variation thereby enhancing a realistic bound for the probability distribution.

The use of Markov inequality to analyze the geometric distribution is inadequate because it gives realistic values for higher values of ' b ' with lower values of ' a ' but becomes unrealistic for other values, hence it cannot be used to establish a proper inequality bound for the probability distribution.

REFERENCES

- [1] Salo-Coste, L (1997). Lectures on finite Markov chains. Lectures on probability theory and statistics (Saint-Flour Summer School, 1996), 301-413, Lecture Notes in Math., 1665, Springer, Berlin.
- [2] Ross, Sheldon (1976) *A First Course in Probability*, Macmillan Publishing Company, New York, New York, 1976.
- [3] Steliga, K and Szydal, D. (2010) "On Markov-type Inequality", International Journal of Pure and Applied Mathematics, 58(2) 137-152.
- [4] Chebyshev, P. (1874). Sur les valeurs limites des integrales, J Math Pure Appl 19: 157-160.
- [5] Feller, W. (1966) An introduction to probability theory and its applications. Vol II, 2nd ed. Wiley, New York pp155.
- [6] Markov, A. (1884). On certain applications of algebraic continued fractions, Ph.D. thesis, St Petersburg.
- [7] Papoulis, A. (1991), Probability, Random Variables, and Stochastic Processes, 3rd ed. McGraw-Hill. ISBN 0-07-100870-5. pp. 113-114.
- [8] Stein, E.M. and Shakarchi, R. (2005) "Real Analysis, Measure Theory, Integration, & Hilbert Spaces", vol. 3, 1st ed., 2005, p.91.
- [9] Clarkson E, Denny JL, Shepp L (2009) ROC and the bounds on tail probabilities via theorems of Dubins and F. Riesz. Ann Appl Prob 19 (1) 467-476 DOI: 10.1214/08-AAP536.
- [10] DasGupta A. (2000). Best constants in Chebyshev inequalities with various applications. Metrika, 51: 185-200.
- [11] McWilliams T.P. (1990). A distribution-free test for symmetry based on a runs statistic. J Am Stat Assoc 85 (412)1130-1133.
- [12] Pitman, Jim. Probability (1993 edition). Springer Publishers. pp. 372.
- [13] Pfeiffer, P. E. and Schum D. A. (1973). *Introduction to Applied Probability*. New York: Academic Press, 1973.

¹DEPARTMENT OF MATHEMATICS AND STATISTICS, FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA, NIGERIA

²DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILORIN, ILORIN, NIGERIA