PROPERTIES OF SOME DISTRIBUTIONS USING CHEBYSHEV'S INEQUALITY APPROACH

Rauf, K.[1]; Oguntolu, F. A.[2]; Isah, A.[2]; Abubakar, U. Y.[2]; & Nafiu, L. A.[2]
[1]Department of Mathematics, University of Ilorin, Ilorin, Nigeria. [2]Department of Mathematics and Statistics, Federal University of Technology, Minna, Nigeria.

Abstract

In this article, we give a simpler proof of Chebyshev inequality and use the result to obtain some properties of Binomial, Poisson and Geometric distributions. Furthermore, analysis of the results has shown that Chebyshev inequality is effective for determining convergence bound of the distributions. Some recent sharpened results are complemented. 2010 Mathematics Subject Classification, 41A50.

Keywords: Chebyshev Inequality, Probability Distributions, Convergence Bound, Measure Space and Sharp.

Introduction

Chebyshev stated, without proof, a theorem called Chebyshev's inequality in 1874. Although it was first formulated without proof by his friend and colleague Irue-jules Bienaym in 1853 (Chebyshev, 1874). However, Markov in 1884 provided a proof in his PhD thesis under the supervision of Chebyshev. It states that in any data sample or probability distribution, nearly all the values are close to the mean value, and provides a quantitative description of nearly all or close to all. We can usually guarantee that more data is a certain number of standard deviations away from the mean if the distribution is clearly known (Steliga.& Szynal, 2010). This inequality is a tool in probability theory; it relates the distribution of numbers in a set (Pitman, 1993). In a general term, the formula helps in determining the number of values that reside in and outside the standard deviation. The standard deviation, however, is a statistically determined number that tells how far away values tend to be from the average of the set. Analytically, about two-thirds of the values should always fall within one standard deviation up or down. It is unlike the empirical relationship between the mean and mode or the rule of thumb that connects the range and standard deviation together.

The inequality is of great significant in the theory of probability distributions and is usually stated for random variables, but can be extended and generalized to a statement about measure spaces (DasGupta, 2000).

In this paper, Chebyshev theorem is proved in a simpler version. The results obtained were used to analysed Binomial, Geometric and Poisson distributions and also to obtain probability bound for some random variables. For recent work see Stein & Shakarchi 2005, Oguntolu 2013, Gauss 1995 and Clarkson et.al. 2009.

Materials and Methods

This section considered the prove of some inequalities and their applications.

The prove of Chebyshev Inequality

Theorem 3.1: Let X be a random variable with mean $E[X]$ and finite variance Var[X]. For any real number $A > 0$, we have

$$\Pr(|X - \mu| \geq A) \leq \frac{Var[X]}{A^2}$$

Proof:

By Chebyshev inequality, $\mu = E[X]$ called the mean (expected value) of the random variable X and

$$\Pr(|X - \mu| \geq A) = \Pr[(X - \mu)^2 \geq A^2]$$

Also, by Markov Inequality we have,

$$\Pr[(X - \mu)^2 \geq A^2] \leq \frac{E(X - \mu)^2}{A^2}$$

$$= \frac{E(X - 2X\mu + \mu^2)}{A^2}$$

$$= \frac{E[X^2] - 2\mu E[X + \mu^2]}{A^2}$$

$$= \frac{E[X^2] - \mu}{A^2}$$

But $Var[X] = E[X^2] - \mu^2$

Hence, $$\Pr[(X - \mu)^2 \geq A^2] \leq \frac{Var[X]}{A^2}$$

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 given by:

$$\Pr[X \geq 0] = \sum_{x=0}^{n} \binom{n}{x} P^x.(1-P)^{n-x}$$

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

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

By Binomial expansion

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

$$= (P + 1 - P)^n = 1$$

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

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

$$= \sum_{x=0}^{n} \frac{n(n-1)!}{(n-x)!(x-1)!} P.P^{x-1} (1-P)^{n-x}$$

$$= nP \sum_{x=0}^{n} \frac{(n-1)!}{(n-x)!(x-1)!} .P^{x-1} (1-P)^{n-x}$$

Hence,

Mean $= E[X] = nP(1) = nP$

Also,

$$E[X^2] = \sum_{x=0}^{n} \frac{x^2 n!}{(n-x)!x!} P^{x} (1-P)^{n-x}$$

$$= \sum_{x=0}^{n} \frac{(x-1)n!}{(n-x)!(x-1)(x-2)!} P^{x} (1-P)^{n-x} + \sum_{x=0}^{n} \frac{n!}{(n-x)!(x-1)!} P^{x} (1-P)^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n(n-1)(n-2)!}{(n-x)!(x-2)!} P.P.P^{x-2} (1-P)^{n-x} + \sum_{x=0}^{n} \frac{n(n-1)!}{(n-x)!(x-1)!} P.P^{x-1} (1-P)^{n-x}$$

$$= n(n-1)P^2 \sum_{x=0}^{n} \frac{(x-n)!}{(n-x)!(x-2)!} P^{x-2} (1-P)^{n-x} + nP \sum_{x=0}^{n} \frac{(n-x)!}{(n-x)!(x-1)!} P^{x-1} (1-P)^{n-x}$$

$$= n(n-1)P^2(1) + nP(1)$$

$$= n(n-1)P^2 + nP$$

Hence,

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

$$= n(n-1)P^2 + nP - (nP)^2$$

$$\Rightarrow nP(1-P)$$

Prove of Chebyshev's Inequality from Binomial distribution

By Binomial distribution, we have

$$\Pr(|X-\mu| \geq t\sigma) = \Pr[(X-\mu)^2 \geq t^2\sigma^2] = \sum_{(x-\mu)^2 = t^2\sigma^2}^{n} \binom{n}{(x-\mu)^2} .P^{(x-\mu)^2} .(1-P)^{n-(x-\mu)^2}$$

$$\leq \frac{1}{(x-\mu)^2} \sum_{(x-\mu)^2 = t^2\sigma^2}^{n} \binom{n}{(x-\mu)^2} .(x-\mu)^2 .P^{(x-\mu)^2} .(1-P)^{n-(x-\mu)^2}$$

$$= \frac{1}{(x-\mu)^2} .E(x-\mu)^2 = \frac{\sigma^2}{(x-\mu)^2}$$

But $\quad (x-\mu)^2 \geq t^2\sigma^2 \Rightarrow \frac{1}{(x-\mu)^2} \leq \frac{1}{t^2\sigma^2}$

$$\le \frac{\sigma^2}{t^2\sigma^2} = \frac{1}{t^2}$$

Hence,

$$\Pr(|X - \mu| \ge t\sigma) \le \frac{1}{t^2}$$

Poisson Distribution

The Poisson distribution is given by:

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

and cummulative probability is

$$\Pr[X \ge 0] = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \;\; = 1$$

and

$$E[X] = \sum_{x=0}^{\infty} \frac{x.\lambda^x e^{-\lambda}}{x!}$$

$$= \lambda e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda}.e^{\lambda} = \lambda$$

Hence,

$$Mean = E[X] = \lambda$$

$$E[X^2] = \sum_{x=0}^{\infty} \frac{x^2.\lambda^x e^{-\lambda}}{x!} = \sum_{x=0}^{\infty} \frac{x.\lambda^x e^{-\lambda}}{(x-1)!}$$

$$= \sum_{x=0}^{\infty} \frac{(x-1).\lambda^x e^{-\lambda}}{(x-1)!} + \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-1)!}$$

$$= \sum_{x=0}^{\infty} \frac{.\lambda^x e^{-\lambda}}{(x-2)!} + \lambda \sum_{x=0}^{\infty} \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!}$$

implying

$$E[X^2] = \lambda^2 e^{-\lambda} \sum_{x=0}^{\infty} \frac{.\lambda^{x-2}}{(x-2)!} + \lambda e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda^2 e^{-\lambda} + \lambda e^{-\lambda}.e^{\lambda} \qquad = \lambda^2 + \lambda$$

Hence, $Variance(\sigma^2) = Var[X] = E[X^2] - (E[X])^2$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

Thus, we have the variance as:

$$\sigma^2 = \lambda$$

The prove of Chebyshev's Inequality from Poisson distribution

By Poisson distribution, we have,

$$\Pr[|X - \mu| \ge t\sigma] \equiv \Pr[(X-\mu)^2 \ge t^2\sigma^2] = \sum_{(x-\mu)^2 = t^2\sigma^2}^{\infty} \frac{\lambda^{(x-\mu)^2} e^{-\lambda}}{(x-\mu)^2!}$$

$$= \frac{(x-\mu)^2}{(x-\mu)^2} \sum_{(x-\mu)^2=t^2\sigma^2}^{\infty} \frac{\lambda^{(x-\mu)^2} e^{-\lambda}}{(x-\mu)^2!}$$

$$\leq \frac{1}{(x-\mu)^2} \sum_{(x-\mu)^2=0}^{\infty} \frac{\lambda.\lambda^{(x-\mu)^2-1} e^{-\lambda}}{[(x-\mu)^2-1]!}$$

Hence,

$$\Pr[|X-\mu| \geq t\sigma] \equiv \Pr[(X-\mu)^2 \geq t^2\sigma^2] = \frac{\lambda}{(x-\mu)^2}.1$$

But,

$$(x-\mu)^2 \geq t^2\sigma^2 \Rightarrow \frac{1}{(x-\mu)^2} \leq \frac{1}{t^2\sigma^2}$$

$$\Pr[|X-\mu| \geq t\sigma] = \Pr[(X-\mu)^2 \geq t^2\sigma^2] \leq \frac{\lambda}{t^2\sigma^2}$$

Since $\sigma^2 = \lambda$

Then, $\Pr[|X-\mu| \geq t\sigma] = \Pr[(X-\mu)^2 \geq t^2\sigma^2] \leq \frac{1}{t^2}$

Geometric Distribution

The Geometric Distribution is given by:

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

and the Cummulative Probability is

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

Let $a = 1-P \Rightarrow a-1 = -P$

then,

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

while

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

and

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

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

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

With the expected value

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

Let $a = 1-P$

Then,

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

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

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

Since $P = 1-a$

Then,

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

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

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

$a^x = (1-P)^x \to 0$ since $0 < (1-P) < 1$

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

$$\Rightarrow \frac{1-P}{P}$$

Hence,

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

with the variance:

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

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

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

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

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

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

Hence,

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

as $x \to \infty$, $a^x \to 0$

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

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

Hence,

$$E[X^2] = \frac{(1-P)^2+(1-P)}{P^2}$$

We have the variance as:

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

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

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

Prove of Chebyshev's Inequality from Geometric distribution

By Geometric distribution we

have $\Pr[|X-\mu|\geq t\sigma] = \Pr[(X-\mu)^2 \geq t^2\sigma^2] = \sum_{(X-\mu)^2=t^2\sigma^2}^{\infty} P(1-P)^{(X-\mu)^2}$

$$= \frac{(X-\mu)^2}{(X-\mu)^2}\sum_{(X-\mu)^2=t^2\sigma^2}^{\infty} P(1-P)^{(X-\mu)^2}$$

$$\leq \frac{1}{(X-\mu)^2}\sum_{(X-\mu)^2=0}^{\infty} P(X-\mu)^2.(1-P)^{(X-\mu)^2}$$

$$= \frac{1}{(X-\mu)^2}\left(E[(X-\mu)^2]\right)$$

But $\mu = E[X]$

$$= \frac{1}{(X-\mu)^2}\left(E[X^2]-[E[X]]^2\right)$$

$$\leq \frac{\sigma^2}{t^2\sigma^2} = \frac{1}{t^2}$$

Hence,

$$\Pr[|X-\mu|\geq t\sigma] \leq \frac{1}{t^2}$$

Tightness of Chebyshev's Inequality

Let us define a random variable X as

$X = \mu + c$ With Probability P

$X = \mu - c$ With Probability P

$X = \mu$ With Probability $1-2P$

$E[X] = P(\mu+c)+P(\mu-c)+\mu(1-2P)$

$E[X] = 2P\mu - 2P\mu + \mu$

$E[X] = \mu$

$Var[X] = E[(X-\mu)^2] = P(\mu + C + \mu)^2 + P(\mu - c - \mu)^2 + (1-2P)(\mu - \mu)^2$

$Var[X] = 2Pc^2$

If we want to find the Probability that the variable deviates from mean by constant c, the bound provided by Chebyshev is

$$\Pr\left(|X - \mu| \geq c\right) \leq \frac{Var[X]}{c^2} = \frac{2Pc^2}{c^2}$$

$$\Pr\left(|X - \mu| \geq c\right) \leq 2P$$

which is sharp (DasGupta (2000)).

Results

Chebyshev inequality is realised from Binomial distribution, Poisson distribution and Geometric distribution, if the initial probability distribution is $P = \frac{1}{a}$ where a is a positive natural number. By plotting the graph of Chebyshev inequalities, we obtain,

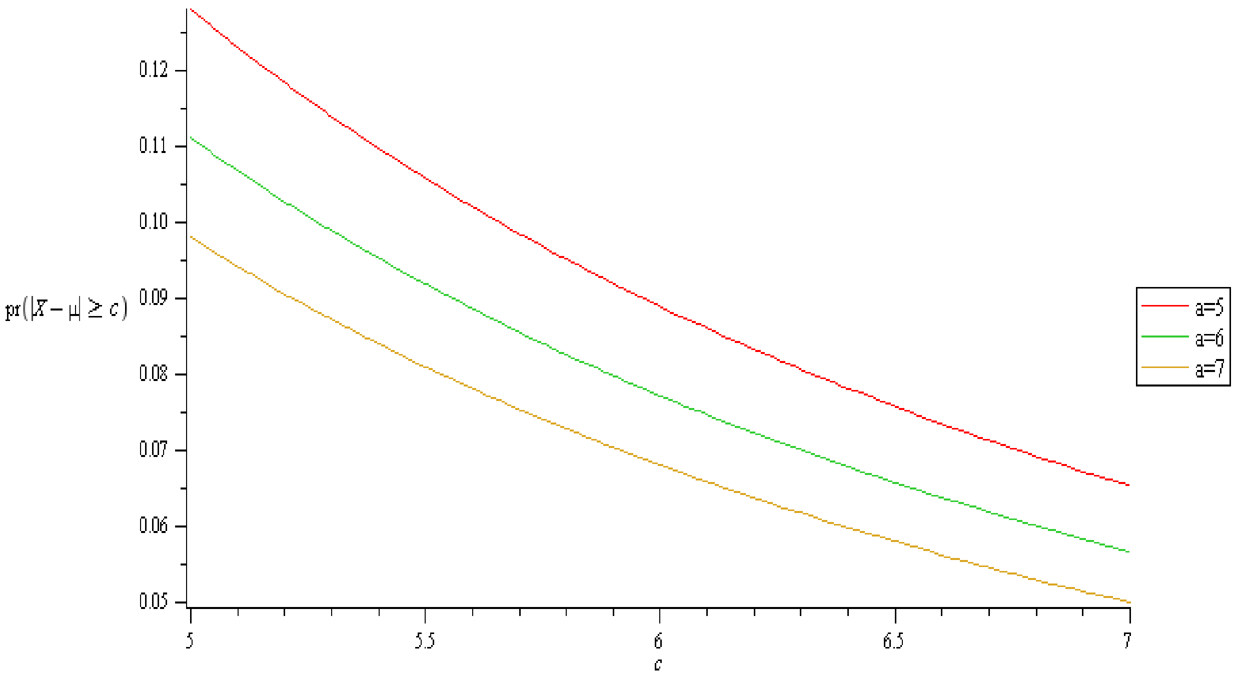
Chebyshev Inequality from binomial distribution

$$\Pr[|X - \mu| \geq t\sigma] \leq \frac{1}{t^2}$$

$$\Pr[|X - \mu| \geq C] \leq \frac{\sigma}{C^2}$$

$$\Pr[|X - n(\frac{1}{a})| \geq t\sqrt{n(\frac{1}{a})(1-\frac{1}{a})}] \leq \frac{1}{t^2} = \frac{\frac{n}{a}(1-\frac{1}{a})}{C^2}$$

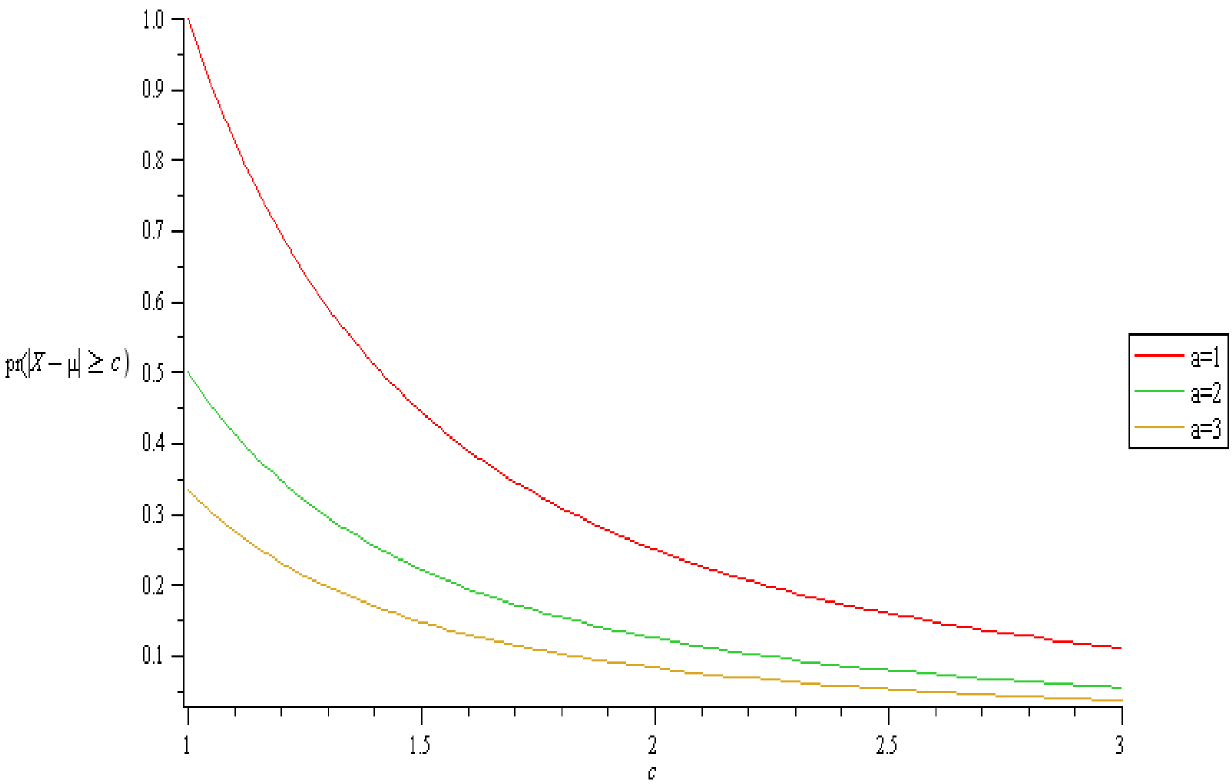
$$\Pr[|X - \frac{n}{a}| \geq \frac{nt(a-1)}{a^2}] \leq \frac{1}{t^2}$$



The graph of Pr[|X − μ| ≥ c] againsts deviation c

Chebyshev Inequality from Poisson distribution

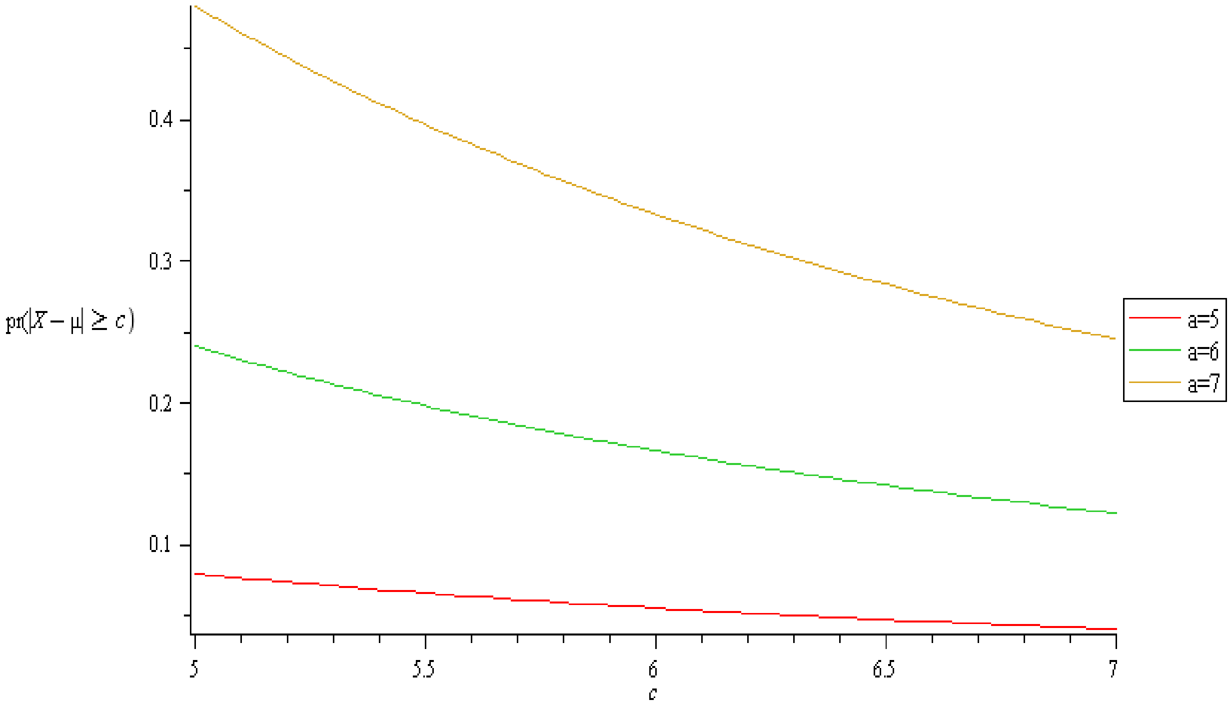
$$\Pr[|\, X - \frac{1}{a} | \geq \frac{t}{a}] \leq \frac{1}{t^2}$$



The graph of Pr[|X − μ| ≥ c] againsts deviation c

Chebyshev Inequality from Geometric distribution

$$\Pr[|X - (a-1)| \geq t.\sqrt{a(a-1)}] \leq \frac{1}{t^2}$$



The graph of Pr[|X − μ| ≥ c] againsts deviation c

Discussion

In Figures 1, 2 and 3, we observed that an increase in the values of $'a'$ will steadily reduce their respective probabilities of the deviation. Also, we observed that as the deviations increase, the probabilities reduce.

Conclusion

The use of Chebyshev inequality to analyze the distribution; Binomial, Poisson or Geometric shows a realistic positive bound for the deviation of the number of trials in the distribution, hence the inequality establishes a good probability bound for certain range of values and as such could be used to forecast any of the distribution.

References

Steliga, K. & Szynal, D. (2010). On Markov-type inequality. *International Journal of Pure and Applied Mathematics*, 58(2) 137-152.

Chebyshev, P. (1874). Sur les valeurs limites des integrals. *Journal of Mathematics, Pure Appl,* 19, 157-160.

Markov, A. (1884). *On certain applications of algebraic continued fractions.* Unpublished Ph.D. thesis, St Petersburg.

Stein, E. M. & Shakarchi, R. (2005). Real analysis, measure theory, integration, and Hilbert spaces, 3, 91.

Clarkson, E., Denny, J. L. & 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.

DasGupta, A. (2000). Best constants in Chebyshev inequalities with various applications. *Metrika,* 51, 185-200.

Gauss, C. F. T. (1995). Combinationis observationum erroribus minimis obnoxiae. Pars Prior. Pars Posterior. Supplementum. Theory of the Combination of Observations Least Subject to Errors. Part One. Part Two. Supplement.Translated by G.W. Stewart. Classics in Applied Mathematics Series, Society for Industrial and Applied Mathematics, Philadelphia.

Pitman, J. (1993). *Probability.* Springer Publishers. pp. 372.

Oguntolu, F. A. (2013). *On some statistical distributions in Markov and Chebyshev inequalities.* Unpublished M. Tech. Dissertation, Federal University of Technology, Minna, Nigeria.