Study on the Expectation of Geometric Distribution

Mohammad Jdeed
PH.D Systems analysis (information processing and control)
Email: dr.m.jdeed@gmail.com

Abstract:

When we want to know the truth of the emergence of a scientific principle, we find in many cases that it is no more than a process of interpretation, description and analysis of existing natural phenomena of various kinds, this is done by identifying the components on which these phenomena depend and the qualities and characteristics that each organism enjoys, which enables us to form a clear strategic vision that helps us in developing pre-solutions to what we expect from the problems resulting from any emergency or exceptional circumstance Most studies in the fields of science need two types of study:

1- Study the real data obtained through observation and description, and then record and tabulate the case information we are studying.
2- the study of phenomena to predict what the phenomenon will become in the near and even distant future,

The statistical analysis puts us in the merits of the current and future situation of the system under study, and statistical analysis is the applied aspect of probability science, through the probability distributions that are used in different fields of science, the probability of the values of the variable can be calculated by applying a mathematical equation called the probability density function and by using distinctive values for these distributions such as expectation, variance and standard deviation From analyzing the data to reach the desired results, after reviewing a number of references, it drew my attention that the expectation of the geometric distribution is calculated using a relationship that varies from one return to another, knowing that these references start from the basic definition of the expectation or from the moment-generating function, and logically the results must be the same, so I prepared this research, through which I explained the reason for the difference, which in turn may have an impact on the results of studies for the systems which works according to this distribution (the importance of this effect depends on the meaning of the expectation for the studied system)

Keywords: Probability; Probability Distributions; Discrete Probability Distributions; Negative Binomial Distribution; Probability Density Function; Moment-Generating Function; Geometric Distribution; Expectation

1. Introduction:

When studying many systems, we find that the course of the study depends on the data collected about the system, where appropriate rules are used for this data, if the data is random and follows a certain probability distribution, then we do not use the distinctive values of this distribution in order to analyze the data, and therefore the relationships used to calculate these values must be accurate relationships and obtained through the application of the basic definition recognized for each of the distinctive values, when we study the Geometric distribution found that some references are used to calculate the expectation of a relationship $E(X) = \frac{q}{p}$, [1,2,3] , and others use a relationship. $E(X) = \frac{1}{p}$, [4 − 14].

In this research, we presented two studies to calculate the expectation of the geometric distribution, which is a special case of the negative binomial distribution, and we explained the reason for the difference and its impact on the results of studies of systems that operate according to the geometric distribution.

Discussion:

We get the expectation using one of two relationships:

1- Starting from the probability density function as follows:

Doi: https://doi.org/10.54216/PAMDA.010104
Received: August 20, 2022 Accepted: December 19, 2022

$$E(X) = \sum_{x \epsilon R_x} x\, f(x) \qquad (1)$$

2- Starting from the moment-generating function as follows:

$$E(X) = \left[(\psi(t))'\right]_{t=0} \qquad (2)$$

Since the geometric distribution is a special case of the negative binary distribution, so in order for the study to be accurate and understandable, we first begin by presenting the negative binomial distribution:

1- Negative binomial distribution (Pascal distribution):[3]

The negative binary distribution (Pascal distribution) is a discrete distribution defined as follows:

Let us have an experience, let it be A an event related to that experience then:

$$P(A) = p \Rightarrow P(A) = 1 - p = q \ \ \& \ \ p + q = 1 \ \ \& \ \ 0 < p < 1$$

Independently we repeat this experiment until we get an appearance of the event a number of times of once we denote B for the experiment in which it is arranged and in which the event will appear a number of times its A magnitude, k this means that in the experiment that we arrange and which we will denote the event will G appear $k + x$ a number of times its magnitude A k , we denote the experiment that will be followed directly by the symbol $x + k - 1$ B in which the A $k - 1$ event will inevitably appear then D it is$A$

$$P(G) = P(D \cap B) = P(B)P(B) \qquad (3)$$

The probability of D a condition B is the probability of the occurrence of A any event

$$P(B) = P(A) = p$$

And the probability B we get from the relationship:

$$P(B) = \binom{x+k-1}{x} p^{k-1} q^x$$

We substitute in the relationship (3) :

$$P(G) = \binom{x+k-1}{x} p^{k-1} q^x\, p$$

$$P(G) = f_X(x) = \binom{x+k-1}{x} p^k q^x \quad ; x = 0,1,2,3,--, k = 1,2,3 \qquad (4)$$

1-1- Moment-generating function of negative binomial distribution (Pascal distribution):

$$\psi_X(t) = \mathrm{E}(e^{tx}) = \sum_{x=0}^{\infty} e^{tx} f_X(x) = \sum_{x=0}^{\infty} e^{tx} \binom{x+k-1}{x} p^k q^x$$

$$\psi_X(t) = \sum_{x=0}^{\infty} e^{tx} \binom{x+k-1}{x} p^k (qe^t)^x =$$

$$\psi_X(t) = p^k \sum_{x=0}^{\infty} \binom{x+k-1}{x} (qe^t)^x = \frac{p^k}{(1-qe^t)^k} \Rightarrow$$

$$\psi_X(t) = \left(\frac{p}{1-qe^t}\right)^k$$

1-2- Expectation of the negative binomial distribution (Pascal distribution) based on the moment-generating function:

$$E(X) = \left[(\psi(t))'\right]_{t=0} = \left[\left(\left(\frac{p}{1-qe^t}\right)^k\right)'\right]_{t=0} = \left[k\left(\frac{p}{1-qe^t}\right)^{k-1}\left(\frac{pqe^t}{(1-qe^t)^2}\right)\right]_{t=0}$$

$$E(X) = k\left(\frac{p}{1-q}\right)^{k-1}\left(\frac{pq}{(1-q)^2}\right) = \frac{kqp}{p^2} = \frac{kq}{p}$$

$$E(X) = \frac{kq}{p}$$

Doi: https://doi.org/10.54216/PAMDA.010104
Received: August 20, 2022 Accepted: December 19, 2022

2. Geometric Distribution:

It is a special case of negative binary distribution when we repeat the experiment until the event appears for the first time, and then stop the experiment This means that, if a random variable expresses failures, then we say that X it is a geometric random variable If X we substitute in the relationship, the k = 1 probability density function of the random variable that follows the geometric distribution becomes as follows: (4)

$$f_X(\mathrm{x}) = \binom{x}{x} pq^x \quad ; x = 0,1,2,---$$

$$P(X = x) = f_X(x) = \binom{x}{x} pq^x = pq^x \quad ; x = 0,1,2,---$$

2-1- The torque-generating function of the geometric distribution from the negative binary distribution:

$$\psi_X(t) = \left(\frac{p}{1 - qe^t}\right)^k \quad k = 1 \rightarrow \psi_X(t) = \frac{p}{1 - qe^t}$$

2-2- Expectation of the geometric distribution based on the negative binary distribution:

$$E(X) = \frac{kq}{p} \xrightarrow{k=1} E(X) = \frac{q}{p} \qquad (5)$$

2- **We calculate the expectation of the geometric distribution from the relationships (1) and (2):**
3-1- Calculation of the expectation of the geometric distribution through relation (1):

Let be a X discrete random variable with a probability density function $f(x) = P(X = x)$ in space R_x

$$P(X = x) = (1 - p)^x p \quad for\ x = 0,1,2,---$$

$$E(X) = \sum\nolimits_{x=0}^{\infty} x(1 - p)^x p \qquad (*)$$

We multiply both sides of the relation $(*)$ by $(1 - p)$

$$(1 - p)E(X) = \sum\nolimits_{x=0}^{\infty} x(1 - p)^{x+1} p$$

Put

$$k = x + 1 \Rightarrow x = k - 1$$

$$(1 - p)E(X) = \sum_{k=1}^{\infty} (k - 1)(1 - p)^k p$$

$$(1 - p)E(X) = \sum\nolimits_{k=1}^{\infty} k(1 - p)^k p - \sum\nolimits_{k=1}^{\infty} (1 - p)^k p$$

$$E(X) - pE(X) = \sum\nolimits_{k=1}^{\infty} k(1 - p)^k p - \sum\nolimits_{k=1}^{\infty} (1 - p)^k p \qquad (**)$$

of the two relationships $(*)$,$(**)$

$$E(X) - pE(X) = E(X) - \sum\nolimits_{k=1}^{\infty} (1 - p)^k p$$

$$pE(X) = \sum\nolimits_{k=1}^{\infty} (1 - p)^k p$$

$$\sum\nolimits_{k=1}^{\infty} (1 - p)^k p = (1 - p) \Rightarrow$$

$$pE(X) = (1 - p) \Rightarrow E(X) = \frac{(1 - p)}{p}$$

$$E(X) = \frac{q}{p}$$

3-4- Calculation of the expectation of the geometric distribution through relation (2):

$$E(X) = \left[\left(\psi(t)\right)'\right]_{t=0} \qquad (2)$$

Doi: https://doi.org/10.54216/PAMDA.010104
Received: August 20, 2022 Accepted: December 19, 2022

$$E(X) = \left[\left(\frac{p}{1-qe^t}\right)'\right]_{t=0} = \left[\frac{pqe^t}{(1-q)^2}\right]_{t=0} = \left[\frac{pqe^t}{p^2}\right]_{t=0} = \frac{q}{p}$$

From the above, we find that the expectation is given by $\boldsymbol{E(X) = \frac{q}{p}}$ **,**

Negative Binomial Distributions: [6]

The probability density function as stated in the reference [6] **takes the following form:**

$$f(x) = \binom{x-1}{k-1} p^k q^{x-k} \quad ; x = k, k+1, k+2, k+3, ---$$

where $x - 1$ represents the number of successes, and we have $k - 1$ successes

4-1- Moment-generating function of negative binomial distribution (Pascal distribution):

$$M_x(t) = \mathrm{E}(e^{tx}) = \sum_{x=k}^{\infty} e^{tx} \binom{x-1}{k-1} p^k q^{x-k}$$

$$M_x(t) = (pe^t)^k \sum_{x=k}^{\infty} \binom{x-1}{k-1} e^{t(x-k)} q^{x-k}$$

If $qe^t < 1$ or $t < -log\,(1-p)$ and the result of the sum is given by $(1-qe^t)^{-k}$ any that

$$M_x(t) = (pe^t)^k (1-qe^t)^{-k} \Rightarrow$$

$$M_x(t) = \left(\frac{pe^t}{(1-qe^t)}\right)^k \; ; t < -log\,(1-p)$$

4-2- Expectation of negative binomial distribution (Pascal distribution) based on the moment-generating function:

$$E(X) = [M_x(t)]'_{t=0}$$

$$E(X) = M'_x(t) = \left[\left(\frac{pe^t}{(1-qe^t)}\right)^k\right]'_{t=0} = \left[\frac{kp^k}{(1-qe^t)^{k+1}}\right]_{t=0}$$

We substitute $\boldsymbol{t = 0}$ **we get the expectation**

$$E(X) = M'_x(0) = \frac{k}{p}$$

3- **Geometric distribution**: It is a special case of negative binary distribution when $k = 1$,Its probability density function becomes as follows:

$$f(x) = \binom{x}{1} pq^{x-1} = p\, q^{x-1} \quad ; x = 1,2,3, ---$$

5-1- The torque-generating function of the geometric distribution based on the negative binary distribution:

$$M_x(t) = \left(\frac{pe^t}{(1-qe^t)}\right)^{\boldsymbol{k}} \xrightarrow{k=1} M_x(t) = \frac{pe^t}{1-(1-p)e^t} \; ; t < -\log\,(1-p)$$

5-2- Expectation of geometric distribution based on negative binary distribution:

$$\boldsymbol{E(X)} = \frac{k}{p} \xrightarrow{k=1} E(X) = \frac{1}{p} \qquad (6)$$

We note that there is a difference between (6) the two relationships and (7) through which we calculate the Expectation of the geometric distribution and the following example shows the difference

Example: In an experiment if the random variable x follows the geometric distribution and the probability of the observed event appearing $\mathrm{p} = \frac{2}{3}$ then the probability of not appearing is $\mathrm{q} = \frac{1}{3}$

The probability density function is according to the first study

Doi: https://doi.org/10.54216/PAMDA.010104
Received: August 20, 2022 Accepted: December 19, 2022

$$f_X(x) = pq^x = \frac{2}{3}\left(\frac{1}{3}\right)^x \quad ; x = 0,1,2,---$$

And the expectation is

$$E(X) = \frac{q}{p} = \frac{3}{2} \times \frac{1}{3} = \frac{1}{2} = 0.5$$

We have the probability density function according to the second study:

$$f(x) = p\, q^{x-1} = \frac{2}{3}\left(\frac{1}{3}\right)^{x-1} \quad ; x = 1,2,3,---$$

And the expectation is

$$E(X) = \frac{1}{p} = \frac{3}{2} = 1.5$$

Conclusion and Results:

The difference in the formula of the expectation between one reference and another reference is due to the fact that some references calculate the expectation from $x = 0$, which means that the event has appeared from the first repetition of the experiment and according to the definition of geometric distribution the experiment is stopped and others take $x = 1$ this means that the event A did not appear the first time of the experiment and must be repeated

We find that it is better to deal with the geometric distribution as a special case of the negative binary distribution because it gives the real value of the expectation, and all users of the geometric distribution must clarify this difference by calculating the expectation.

References:

[1] Michael J. Evans and Jeffrey S. Rosenthal, Probability and Statistics, the Science of Uncertainty, Second Edition, University of Toronto, 2009
[2] JAYL. DEVORE,Probability and Statistics for Engineering and the sciences,California Polytechnic Stste University,2010
[3] Ahmed Al-Wasouf, Mohammad jdeed, Kausar Nazeha, Introduction to statistics and probability, Tishrieen University,2008
[4] Grinstead and Snell's Introduction to Probability, The CHANCE Project1 ,Version dated 4 July 2006
[5] Azzam Sabry, Statistical analysis between theory and practice, Dar Almanhajiah Publishing ,Amman,2015
[6] Abdu Al jabbar Mudhi ,Introduction to Probability,
[7] Theory,DarAlmassira, Amman, 2011
[8] A Short Introduction to Probability, D.P. Kroese, 2018
[9] Sheldon M. Ross, Introduction to Probability Models, University of Southern California Los Angeles, California, Elsevier Inc, 2014Now
[10] John Wiley & Sons Ltd, Fundamentals of Probability and Statistics for Engineers, , New York, USA,2004
[11] Prasanna Sahoo, Probability and Mathematical Statistics, Louisville,USA,2015

Doi: https://doi.org/10.54216/PAMDA.010104
Received: August 20, 2022 Accepted: December 19, 2022