Aarhus university

Computer-Science

Introduction to probability and statistics

Handin 4

Author
Søren M. Damsgaard

Student number
202309814

September 25, 2025

Let $0 < p < 1$ and $X \approx$ Geometric(p).

Calculate The cumulative distribution function (CDF) $F_X(x) = P(X \leq x)$ for all $x = 1, 2, 3, ...$.

Show that X has the 'lack of memory' property, i.e.

$$P(X > m + k | X > m) = P(X > k) \text{ for all } k, m = 1, 2, 3, ... \tag{1}$$

Finding the CDF (Dora's solution)

The CDF is essentially the sum of all the probabilities up to and including x for discrete variables.
This means that we can express it as:

$$F_X(x) = P(X \leq x) = \sum_{i=1}^{x} P(X = i) \quad \text{for } x = 1, 2, 3, ... \tag{2}$$

From Def 3.5 [1] we know that the PMF of a geometric distribution is:

$$P(X = i) = (1 - p)^{i-1} p \tag{3}$$

This means that we can rewrite the CDF as:

$$F_X(x) = \sum_{i=1}^{x} (1 - p)^{i-1} p \tag{4}$$

Move the constant out of the summation:

$$F_X(x) = p \sum_{i=1}^{x} (1 - p)^{i-1} \tag{5}$$

We will now show that the current summation is a geometric series.
A geometric series is defined as:

$$a + ax + ax^2 + ax^3 + ... + ax^{n-1} = \sum_{k=0}^{n-1} ax^k = a\frac{1 - x^n}{1 - x} \tag{6}$$

from example 1.12 equation 1.3 [1].
We can see that our summation is a geometric series with $a = 1$, $x = (1 - p)$ and $n = x$.
This means that we can rewrite the CDF as:

$$F_X(x) = p \cdot \frac{1 - (1 - p)^x}{1 - (1 - p)} \tag{7}$$

which simplifies to:

$$F_X(x) = 1 - (1-p)^x \tag{8}$$

Note that it's important when we use the geometric series we, it has to start at indexx 0, so we had to re

Show X has dementia (lack of memory property)

Reiterating equation(1)

$$P(X > m+k|X > m) = P(X > k) \text{ for all } k, m = 1, 2, 3, ...$$

We use the definition of conditional probability from chapter 1.4 [1] to do legal crime:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{9}$$

We can then rewrite the right side of equation(1) as:

$$P(X > m+k|X > m) = \frac{P((X > m+k) \cap (X > m))}{P(X > m)} = \frac{P(X > m+k)}{P(X > m)} \tag{10}$$

We note here that $P((X > m+k) \cap (X > m)) = P(X > m+k)$ since $X > m+k \implies X > m$.

We can now use the CDF we found earlier to express $P(X > x)$ using the complement, like from example 3.10 [1]:

$$P(X > x) = 1 - P(X \leq x) = 1 - F_X(x) \tag{11}$$

We can then take equation(10) and throw our new math at it like Kobe Bryant on January 22, 2006 (60 points on 50 shots):

$$P(X > m+k|X > m) = \frac{1 - F_X(m+k)}{1 - F_X(m)} \tag{12}$$

We can then insert the value of the CDF:

$$P(X > m+k|X > m) = \frac{1 - (1 - (1-p)^{m+k})}{1 - (1 - (1-p)^m)} \tag{13}$$

which we can beautify a little bit by removing the negative vibes:

$$P(X > m+k|X > m) = \frac{(1-p)^{m+k}}{(1-p)^m} \tag{14}$$

And as a finishing touch we will put the crown on it by simplifying the fraction even further:

$$P(X > m+k|X > m) = (1-p)^k \tag{15}$$

Now we can see that the right side of equation(1) is:

$$P(X > k) = 1 - F_X(k) = 1 - (1 - (1-p)^k) = (1-p)^k \tag{16}$$

References

[1] Hossein Pishro-Nik. *Introduction to Probability ,Statistics and Random Processes.* Kappa Research, LLC, 2014.