Lecture 03: Transformations and Expectations

Mathematical Statistics I, MATH 60061/70061

Tuesday September 7, 2021

Reference: Casella & Berger, 2.1-2.2

Location-scale transformation

Let X be a random variable and $Y = \sigma X + \mu$, where σ and μ are constant with $\sigma > 0$. Then we say that Y has been obtained as a location-scale transformation of X.

Using location-scale transformations to study continuous R.V.s:

- Start with the simplest form of the same family
- Extend to general cases using location-scale transformations

For example, we could figure out the mean and variance of the $\mathrm{Unif}(0, 1)$ distribution first, and then extend to $\mathrm{Unif}(a, b)$.

The location-scale strategy for Uniform distributions

Let $U \sim \mathrm{Unif}(0, 1)$

$$E(U) = \int_0^1 x dx = \frac{1}{2},$$
$$E(U^2) = \int_0^1 x^2 dx = \frac{1}{3},$$
$$\mathrm{Var}(U) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

The R.V. $\tilde{U} = a + (b - a)U$ is distributed $\mathrm{Unif}(a, b)$. By linearity,

$$E(\tilde{U}) = E(a + (b - a)U) = a + (b - a)E(U) = \frac{a + b}{2}.$$

Using the properties of variance,

$$\mathrm{Var}(\tilde{U}) = \mathrm{Var}(a + (b - a)U) = (b - a)^2 \mathrm{Var}(U) = \frac{(b - a)^2}{12}.$$

PDF of a location-scale transformation

Let X have PDF f_X, and let $Y = a + bX$, with $b \neq 0$. Let $y = a + bx$, to mirror the relationship between Y and X. Then $\frac{dy}{dx} = b$, so the PDF of Y is

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right| = f_X\left(\frac{y-a}{b}\right)\frac{1}{|b|}.$$

Geometric distribution

Consider a sequence of independent Bernoulli trials, each with the same success probability $p \in (0, 1)$, with trials performed until a success occurs. Let X be the number of *failures* before the first successful trial. Then X has the **Geometric distribution** with parameter p, $X \sim \mathrm{Geom}(p)$.

Geometric distribution

Consider a sequence of independent Bernoulli trials, each with the same success probability $p \in (0, 1)$, with trials performed until a success occurs. Let X be the number of *failures* before the first successful trial. Then X has the **Geometric distribution** with parameter p, $X \sim \mathrm{Geom}(p)$.

The PMF of X is

$$P(X = k) = q^k p,$$

for $k = 0, 1, 2, \ldots$, where $q = 1 - p$.

Geometric distribution

Consider a sequence of independent Bernoulli trials, each with the same success probability $p \in (0, 1)$, with trials performed until a success occurs. Let X be the number of *failures* before the first successful trial. Then X has the **Geometric distribution** with parameter p, $X \sim \mathrm{Geom}(p)$.

The PMF of X is

$$P(X = k) = q^k p,$$

for $k = 0, 1, 2, \ldots$, where $q = 1 - p$.

This is a valid PMF (it sums to 1):

$$\sum_{k=0}^{\infty} q^k p = p \sum_{k=0}^{\infty} q^k = p \cdot \frac{1}{1 - q} = 1.$$

Geometric expectation

Let $X \sim \mathrm{Geom}(p)$. By definition,

$$E(X) = \sum_{k=0}^{\infty} kq^k p,$$

where $q = 1 - p$.

- It is not a geometric series because of the extra k.
- Each term looks similar to kq^{k-1}, the derivative of q^k.

Geometric expectation

Let $X \sim \mathrm{Geom}(p)$. By definition,

$$E(X) = \sum_{k=0}^{\infty} kq^k p,$$

where $q = 1 - p$.

- It is not a geometric series because of the extra k.
- Each term looks similar to kq^{k-1}, the derivative of q^k.

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

Geometric expectation

Let $X \sim \mathrm{Geom}(p)$. By definition,

$$E(X) = \sum_{k=0}^{\infty} k q^k p,$$

where $q = 1 - p$.

- It is not a geometric series because of the extra k.
- Each term looks similar to kq^{k-1}, the derivative of q^k.

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

$$\Rightarrow \sum_{k=0}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$

Geometric expectation

Let $X \sim \mathrm{Geom}(p)$. By definition,

$$E(X) = \sum_{k=0}^{\infty} kq^k p,$$

where $q = 1 - p$.

- It is not a geometric series because of the extra k.
- Each term looks similar to kq^{k-1}, the derivative of q^k.

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

$$\Rightarrow \sum_{k=0}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2}$$

$$\Rightarrow E(X) = \sum_{k=0}^{\infty} kq^k p = pq\frac{1}{(1-q)^2} = \frac{q}{p}.$$

Negative Binomial distribution

In a sequence of independent Bernoulli trials with success probability $p \in (0, 1)$, if X is the number of *failures* before the rth success, then X is said to have the **Negative Binomial distribution** with parameters r and p, $X \sim \mathrm{NBin}(r, p)$.

Negative Binomial distribution

In a sequence of independent Bernoulli trials with success probability $p \in (0,1)$, if X is the number of *failures* before the rth success, then X is said to have the **Negative Binomial distribution** with parameters r and p, $X \sim \mathrm{NBin}(r,p)$.

The PMF of X is

$$P(X = n) = \binom{n+r-1}{r-1} p^r q^n,$$

for $n = 0, 1, 2, \ldots$, where $q = 1 - p$.

Negative Binomial distribution

In a sequence of independent Bernoulli trials with success probability $p \in (0, 1)$, if X is the number of *failures* before the rth success, then X is said to have the **Negative Binomial distribution** with parameters r and p, $X \sim \mathrm{NBin}(r, p)$.

The PMF of X is

$$P(X = n) = \binom{n+r-1}{r-1} p^r q^n,$$

for $n = 0, 1, 2, \ldots$, where $q = 1 - p$.

- Imagine a string of 0's and 1's, with 1's representing successes.
- The probability of any specific string of n 0's and r 1's is $p^r q^n$.
- The string terminate in the rth 1; there are $(r-1)$ 1's in the first $n+r-1$ positions. So there are $\binom{n+r-1}{r-1}$ such strings.

Negative Binomial expectation

A Negative Binomial r.v. $X \sim \mathrm{NBin}(r, p)$ can be viewed as the number of failures before the rth success in a sequence of independent Bernoulli trials with success probability p.

- X_1: the # of failures until the first success, $X_1 \sim \mathrm{Geom}(p)$
- X_2: the # of failures between the first success and the second success, $X_2 \sim \mathrm{Geom}(p)$
- $\ldots$
- X_r: the # of failures between the $(r-1)$th success and the rth success, $X_r \sim \mathrm{Geom}(p)$

So a Negative Binomial r.v. can be represented as a sum of i.i.d. Geometrics $X = X_1 + \cdots + X_r$, where the X_i are i.i.d. $\mathrm{Geom}(p)$.

Negative Binomial expectation

A Negative Binomial r.v. $X \sim \mathrm{NBin}(r, p)$ can be viewed as the number of failures before the rth success in a sequence of independent Bernoulli trials with success probability p.

- X_1: the # of failures until the first success, $X_1 \sim \mathrm{Geom}(p)$
- X_2: the # of failures between the first success and the second success, $X_2 \sim \mathrm{Geom}(p)$
- $\ldots$
- X_r: the # of failures between the $(r-1)$th success and the rth success, $X_r \sim \mathrm{Geom}(p)$

So a Negative Binomial r.v. can be represented as a sum of i.i.d. Geometrics $X = X_1 + \cdots + X_r$, where the X_i are i.i.d. $\mathrm{Geom}(p)$.

By linearity, the expected value of X is

$$E(X) = E(X_1) + \cdots + E(X_r) = r \cdot \frac{q}{p}.$$