

Week 5 – Expectation, variance, and transformations of RVs

Juan Pablo Lewinger

Last class

- Continuous random variables:
 - Take infinite uncountable values
 - Have a probability density function (pdf)
- Continuous RV are completely characterized by their pdf or cdf
- Uniform, $U[a, b]$
- Exponential, $Exp[\lambda]$
- Normal, $N(\mu, \sigma)$
- Pareto, $Pareto(x_m, \alpha)$
- Mixtures of distributions
 - Discrete + Discrete = Discrete
 - Continuous + Continuous = Continuous
 - Discrete + Continuous = **Neither discrete nor continuous**

Last class

For a continuous random variable, it follows from the definition of pdf and the fundamental theorem of calculus that for $a \leq b$:

$$P(a < X \leq b) = \int_a^b f_X(t) dt = F_X(x) \Big|_a^b = F(b) - F(a)$$

But $P(a < X \leq b) = F(b) - F(a)$ is true for any random variable (discrete, continuous or mixed):

$A = \{X \leq a\}, B = \{X \leq b\}$ Clearly, $A \subset B$ and $B - A = \{a < X \leq b\}$

$$P(a < X \leq b) = P(B - A) = P(B) - P(A) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

For a continuous RV (but not for a discrete or mixed) this also equals $P(a \leq X \leq b), P(a < X < b), \dots$

Expectation for discrete RVs

The expected value is a weighted average of the values a random variable takes, weighted by the probability of taking those values. It's the center of 'gravity' where the distribution 'balances'.

For a **Discrete** random variable X , $f_X(x)$ the probability mass function of X , and $\{x_1, x_2, \dots\}$ is the support of X .

$$E[X] = \sum_{x_i \in \text{supp}(X)} x_i f_X(x_i)$$

The support of a discrete random variable X is the set of points that X takes with non-zero probability:
 $\text{supp}(X) = \{x_i : P(X = x_i) > 0\}$

Expectation discrete examples: Bernoulli

Bernoulli trial, $X \sim \text{Bernoulli}(p)$

$$X = \begin{cases} 1 & p \\ 0 & 1 - p \end{cases}$$

$$f_X(0) = 1 - p,$$

$$f_X(1) = p \quad (0 \leq p \leq 1)$$

$$E[X] = 0 \times (1 - p) + 1 \times p = p$$

Expectation discrete examples: Binomial

$X \sim \text{Binomial}(n, p)$; models the number of successes in n trials

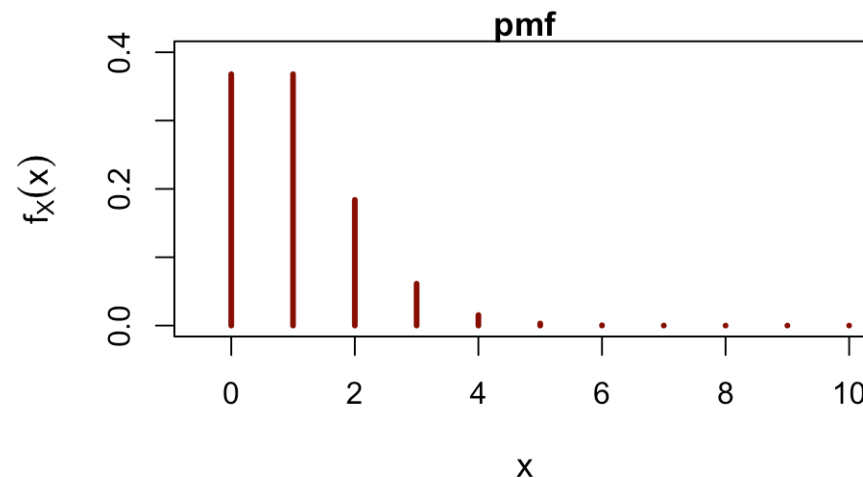
$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n$$

$$\begin{aligned} E[X] &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} = \\ &= \sum_{k=1}^n k \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} = np \sum_{k=1}^n \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} (1-p)^{n-k} = \\ &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = \\ &= np(p + (1-p))^{n-1} = np \end{aligned}$$

Expectation discrete examples: Poisson

A discrete random variable X is said to have a $Poisson(\lambda)$ distribution with parameter $\lambda > 0$ if $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$ for $k = 0, 1, 2, \dots$

Used to model number of events happening in a period of time e.g. number of mutations per unit length in a DNA strand, number of new patients (incidence rates), number of phone calls/particles arriving in a system, etc.



Expectation examples: Poisson (contd)

$$\begin{aligned} E[X] &= \sum_{k=0}^{\infty} k f_X(k) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \\ &= \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

Expectation for continuous RVs

For a **Continuous** random variable X , with $f_X(x)$ the probability density function of X , the expectation is defined as:

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

NOTE: Expectation may not exist E.g. Cauchy distribution $f(x) = \frac{1}{\pi(1+x^2)}$,
 $-\infty < x < +\infty$

Expectation examples: continuous with finite support

$$X \sim F(x)$$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 2x^2 - x^4 & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

$$f(x) = F'(x) = \begin{cases} 4x - 4x^3 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^1 x(4x - 4x^3) dx = 4 \left(\frac{x^3}{3} - \frac{x^5}{5} \right) \Big|_0^1 = \frac{8}{15}$$

Expectation examples: continuous with infinite support

$$X \sim \text{Exp}[\lambda], f(x) = \lambda e^{-\lambda x} \text{ for } x \geq 0$$

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \int_0^{+\infty} t e^{-t} dx$$

(Using $t = \lambda x, dt = \lambda dx$)

$$E[X] = \frac{1}{\lambda} \int_0^{+\infty} t e^{-t} dx = \frac{1}{\lambda} \left(-t e^{-t} \Big|_0^{+\infty} - \int_0^{+\infty} e^{-t} dt \right) = \frac{1}{\lambda} (-e^{-t}) \Big|_0^{+\infty} = \frac{1}{\lambda}$$

(integrating by parts)

Expectation for a mixed random variable

For a **mixed** random variable X with cdf $F(x) = pF_1(x) + (1 - p)F_2(x)$, with F_1 continuous and F_2 discrete:

$$E[X] = p \int_{-\infty}^{+\infty} x f_1(x) dx + (1 - p) \sum_{x_i} x_i f_2(x_i)$$

where $f_1(x)$ is the density for the continuous component, $f_1(x) = F_1'(x)$, and $f_2(x)$ is the mass function for the discrete component.

Probabilities as expectations

For a set $A \subset \Omega$, the random variable $I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$

is called the indicator function of the set A .

What is the distribution of I_A ?

$$I_A \sim \text{Bernoulli}(p), p = P(A) \implies E[I_A] = P(A)$$

Allows us to work with random variables (indicator functions) instead of sets and expectations instead of probabilities

Exercise: If $A, B \subset \Omega$, what are $I_A I_B$, $\min(I_A, I_B)$, $\max(I_A, I_B)$?

Transformations of random variables

Example: $X \sim F(x)$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 2x^2 - x^4 & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

What is the cdf of $Y = -\sqrt{X} + 1$?

First, $0 \leq X \leq 1 \iff -\sqrt{2} \leq -\sqrt{X} + 1 \leq -1$

Transformations of random variables

For $-\sqrt{2} \leq y \leq -1$

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = \\ &= P(-\sqrt{X+1} \leq y) \\ &= P(-y \leq \sqrt{X+1}) = \\ &= P(y^2 - 1 \leq X) = \\ &= 1 - P(X < y^2 - 1) = 1 - P(X \leq y^2 - 1) \\ &= 1 - F_X(y^2 - 1) = 1 - 2(y^2 - 1)^2 + (y^2 - 1)^4 \end{aligned}$$

$$F_Y(y) = \begin{cases} 0 & \text{if } y < -1 \\ 1 - 2(y^2 - 1)^2 + (y^2 - 1)^4 & \text{if } -\sqrt{2} \leq y \leq -1 \\ 1 & \text{if } y > -\sqrt{2} \end{cases}$$

Change-of-variable formula for the expectation

Continuous X :

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

Discrete X :

$$E[g(X)] = \sum_{x_i} g(x_i) f_X(x_i)$$

Allows us to compute the expectation of $Y = g(X)$ without deriving the pdf or pmf of $g(X)$!!

Change-of-variable formula example

- In the previous example computing $E[Y] = \int_{-\sqrt{2}}^{-1} y f_Y(y) dy$ seems to require knowing/deriving $f_Y(y) = F_Y'(y)$ and integrating using the pdf of Y (a big mess)
- The change-of-variable formula allows us to use a shortcut:
- $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$
- Only require us to integrate using the pdf of X !

$$E[-\sqrt{X} + 1] = \int_0^1 -(\sqrt{x} + 1)(2x^2 - x^4) dx$$

- Using Mathematica: `Integrate[-Sqrt[(1 + x)](2x^2 - x^4), {x, 0, 1}]`

$$E[-\sqrt{X} + 1] = -\frac{4}{693}(103\sqrt{2} - 40)$$

Variance

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

$\text{Var}[X]$ is the average squared deviation from the mean. Measure of dispersion/concentration.

Continuous X

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f_X(x) dx$$

Discrete X

$$E[X^2] = \sum_{x_i} x_i^2 f_X(x_i)$$

Next Week

- Read *IPS 5.1.0-5.1.4 and 5.2.0-5.2.3* (do the exercises!)