CSCI 3022: Intro to Data Science
Lecture 10: Expectation of Discrete and Continuous Random Variables

Rachel Cox

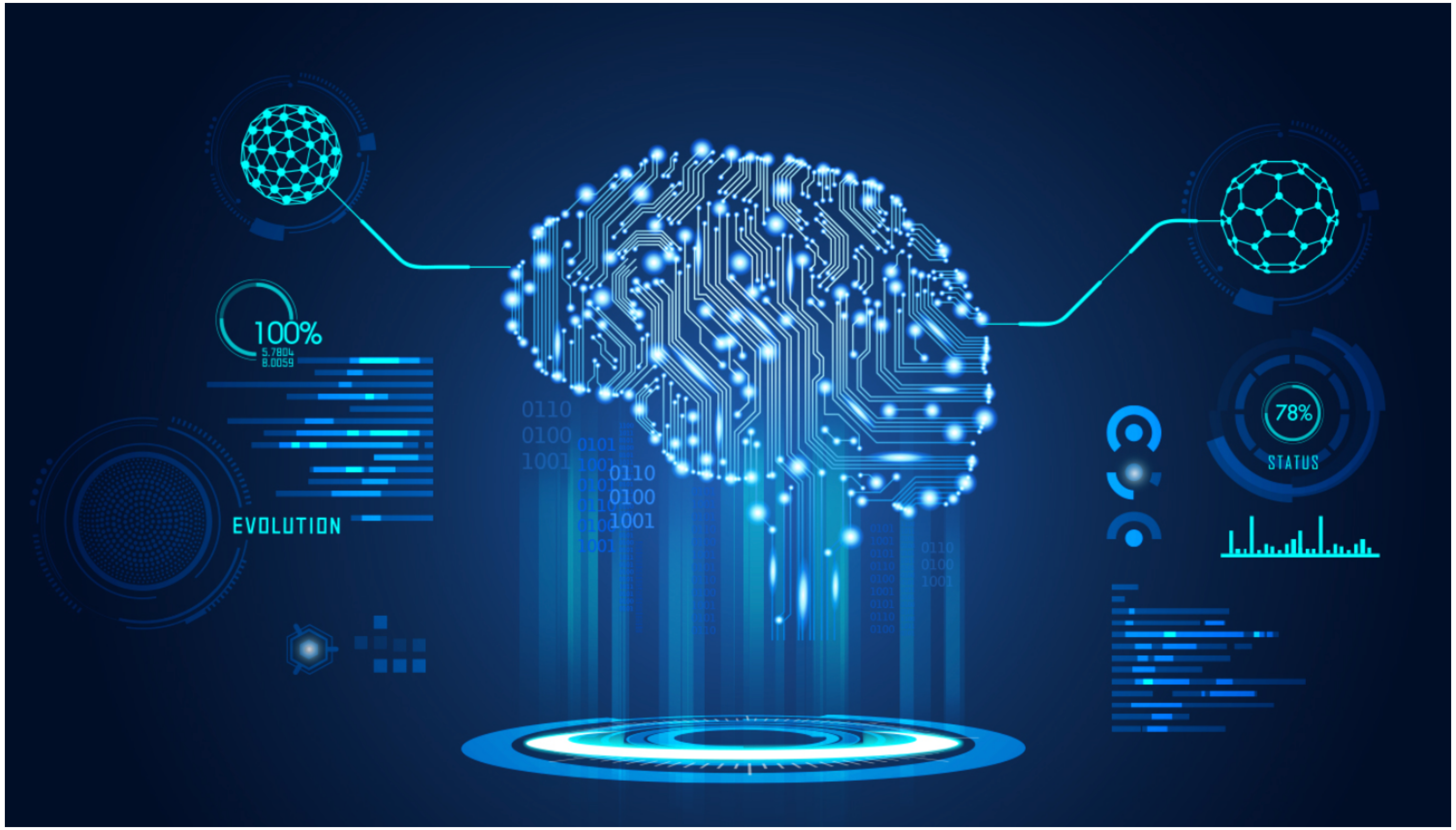
Department of Computer Science

Announcements & Reminders

What will we learn today?

- ❑ Expected Value
- ❑ Expectation of Random Variables
- ❑ Expectation of Functions of Random Variables
- ❑ Linearity of Expectation
- ❑ Expected Values and their relationship to the pmf/pdf

- ❑ *A Modern Introduction to Probability and Statistics, Chapter 7*



Review from Last Time

A **probability mass function** (pmf) is the map between the random variable's values and the probabilities of those values.

$$f(a) \quad = \quad P(X = a)$$

A random variable X is **continuous** if for some function $f: \mathbb{R} \to \mathbb{R}$ and for any numbers a and b with $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

The **probability density function** (pdf) must satisfy:

1) $f(x) \geq 0$ for all x
2) $\int_{-\infty}^{\infty} f(x)\, dx \;\; = 1$

Expected Value: Discrete Random Variables

The **expectation** or **expected value** of a discrete random variable X that takes the values $a_1, a_2, \ldots.$ and with pmf p is given by

$$E[X] = \sum_i a_i P(X = a_i) = \sum_i a_i p(a_i)$$

Intuition: Think of masses of weight $p(a_i)$ placed at the points a_i, $E[X]$ is the balance point.

Center of Mass (from Calculus/Physics) is an analogous concept.

Example: Suppose that I write the homework questions as either: easy (takes 10 minutes), medium (60 minutes), or hard (120 minutes). The probability that each question is easy/medium/hard is: 0.2, 0.3, 0.5 respectively.

If a homework consists of 5 questions, what's the average time it takes to do the homework?

Expected Value: Discrete Random Variables

The **expectation** or **expected value** of a discrete random variable X that takes the values $a_1, a_2, \ldots.$ and with pmf p is given by

$$E[X] = \sum_i a_i P(X = a_i) = \sum_i a_i p(a_i)$$

Example: Suppose that I write the homework questions as either: easy (takes 10 minutes), medium (60 minutes), or hard (120 minutes). The probability that each question is easy/medium/hard is: 0.2, 0.3, 0.5 respectively.

If a homework consists of 5 questions, what's the average time it takes to do the homework?

easy = 10 minutes
medium = 60 minutes
hard = 120 minutes

p(easy) = 0.2
p(medium) = 0.3
p(hard) = 0.5

$$[10(0.2) + 60(0.3) + 120(0.5)] * 5 = [2 + 18 + 60] * 5 = \boxed{400 \text{ minutes}}$$

Expected Value: Discrete Random Variables

Example: Let X be a Bernoulli random variable with parameter p. What is $E[X]$?

$$f(x) = \begin{cases} p & x=1 \\ 1-p & x=0 \end{cases}$$

$$E[X] = \sum_i a_i \, p(a_i) = 1 \cdot f(1) + 0 \cdot f(0)$$

$$= 1 \cdot p + 0 \cdot (1-p)$$

$$= p$$

$$\boxed{E[X] = p}$$

Expected Value: Discrete Random Variables

Example: Suppose you and a friend are avoiding studying by each rolling a fair die. You decide that the first time that you roll the same number, you'll go back to work.

What is the expected number of times you'll roll the dice before getting a match?

This activity is modeled by a geometric distribution.

$$f(k) = (1-p)^{k-1} p$$

$$E[x] = \sum_i a_i f(a_i)$$

$$= \sum_{i=1}^{\infty} i (1-p)^{i-1} \cdot p$$

Recall that $\frac{d}{dx} x^i = i x^{i-1}$

Notice:

$$\frac{d}{dp} \sum_{i=1}^{\infty} (1-p)^i = \sum_{i=1}^{\infty} \frac{d}{dp} (1-p)^i$$

$$= \sum_{i=1}^{\infty} i (1-p)^{i-1} (-1)$$

$$-\frac{d}{dp} \sum_{i=1}^{\infty} (1-p)^i = \sum_{i=1}^{\infty} i (1-p)^{i-1}$$

Expected Value: Discrete Random Variables

geometric distribution

Example: Suppose you and a friend are avoiding studying by each rolling a fair die. You decide that the first time that you roll the same number, you'll go back to work.

What is the expected number of times you'll roll the dice before getting a match?

$$E[x] = p \underbrace{\sum_{i=1}^{\infty} i(1-p)^{i-1}}$$

$$= p\left(-\frac{d}{dp} \underbrace{\sum_{i=1}^{\infty} (1-p)^{i}}_{\text{infinite geometric series}}\right) \qquad \frac{a}{1-r}$$

$$= p \cdot -\frac{d}{dp}\left(\frac{1-p}{1-(1-p)}\right)$$

$$= p \cdot -\frac{d}{dp}\left(\frac{1-p}{p}\right)$$

$$= -p \cdot \left(\frac{p(-1)-(1-p)\cdot 1}{p^2}\right)$$

$$= -p \cdot \left(\frac{-p-1+p}{p^2}\right)$$

$$= -p \cdot \left(-\frac{1}{p^2}\right)$$

$$= \frac{1}{p}$$

$$\boxed{E[x] = 1/p}$$

Expected Value: Continuous Random Variables

How does the expected value change from the discrete case to the continuous one?

discrete: $E[X] = \sum_i a_i \, p(a_i)$

continuous: $\boxed{E[X] = \int_{-\infty}^{\infty} x \, f(x) \, dx}$

Expected Value: Continuous Random Variables

The **expectation** or **expected value** of a continuous random variable X with probability density function f is:

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

Intuition: Think of a single big rock balancing on a fulcrum.

Kummakivi rock in Finland

Expected Value: Continuous Random Variables

IBP: $\int u\,dv = uv - \int v\,du$

Example: The lifetime (in years) of a certain brand of battery is Exponentially distributed with parameter $\lambda = 0.25$.

1) How long, on average, will this battery last?
2) What are the units of λ?

1) The pdf for the exponential distribution: $f_x(x) = \lambda e^{-\lambda x}$

$$E[x] = \int_{-\infty}^{\infty} x \lambda e^{-\lambda x}\,dx$$

$$= \lambda \int_{0}^{\infty} x e^{-\lambda x}\,dx$$

Integration by parts

$u = x \qquad v = -\frac{1}{\lambda} e^{-\lambda x}$

$du = 1\,dx \qquad dv = e^{-\lambda x}\,dx$

$$= \lambda \left[x e^{-\lambda x} \Big|_0^{\infty} - \int_0^{\infty} -\frac{1}{\lambda} e^{-\lambda x}\,dx \right]$$

$$= \lambda \left[x e^{-\lambda x} \Big|_0^{\infty} + \frac{1}{\lambda}\left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^{\infty} \right]$$

Expected Value: Continuous Random Variables

Example: The lifetime (in years) of a certain brand of battery is Exponentially distributed with parameter $\lambda = 0.25$.

1) How long, on average, will this battery last?
2) What are the units of λ? deaths/year

1) (continued)

$$E[x] = \lambda\left[xe^{-\lambda x}\Big|_0^\infty - \frac{1}{\lambda^2} e^{-\lambda x}\Big|_0^\infty \right]$$

$$= \lim_{t\to\infty} \lambda x e^{-\lambda x}\Big|_0^t - \frac{1}{\lambda} e^{-\lambda x}\Big|_0^t$$

$$= \lim_{t\to\infty}\left[\underbrace{\cancel{\lambda t e^{-\lambda t}}}_{\to 0} - 0 - \left(\underbrace{\cancel{\frac{1}{\lambda} e^{-\lambda t}}}_{\to 0} - \frac{1}{\lambda} e^{0} \right)\right]$$

$$= \frac{1}{\lambda}$$

$$\boxed{\text{For our problem } E[x] = 4 \text{ years!}}$$

$\Rightarrow$ $\boxed{\text{For the exponential distribution } E[x] = 1/\lambda}$

Expected Value: Continuous Random Variables

Example: Suppose you have observed that, on average, 300 cars cross a particular bridge every hour. How much time do you expect to wait between two cars crossing the bridge?

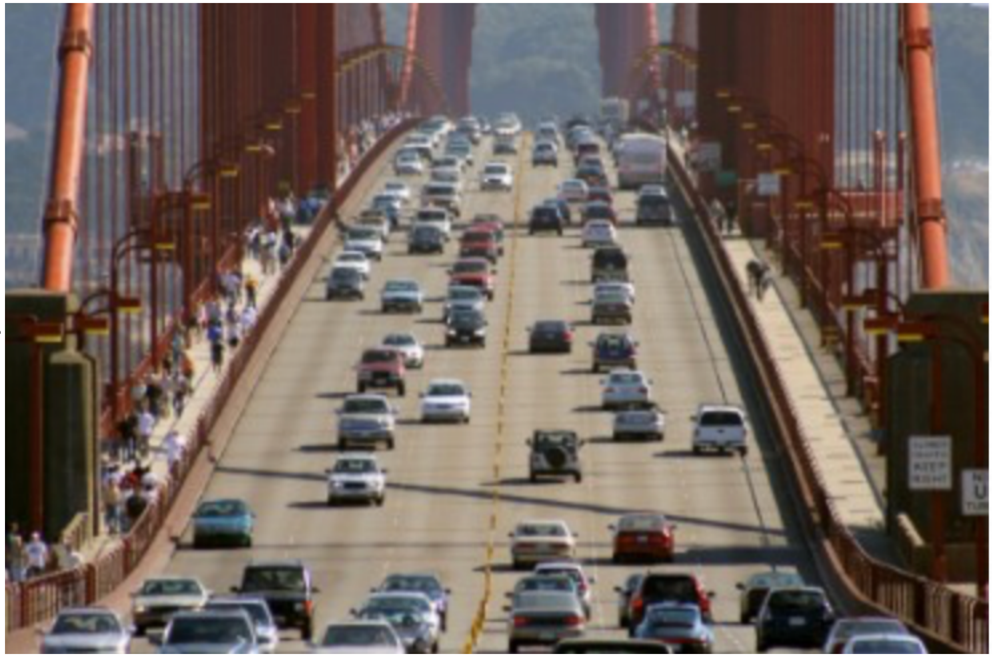
Car "arrivals" is a Poisson process. $\lambda = 300$

$\Rightarrow$ Exponential distribution deals with the wait time between arrivals.

$$E[X] = \frac{1}{300 \text{ car/hour}} = \frac{1 \text{ hour}}{300 \text{ cars}}$$

$$= \frac{1}{300} \text{ hour/car}$$

$$= 12 \text{ seconds/car}.$$



Expectation of Functions of Random Variables

Often we want to compute the expectation of a function of a random variable, instead of the random variable itself.

e.g. $E[X^2]$ instead of $E[X]$

Example: Suppose an architect is designing a community and wants to maximize the diversity in the size of his square buildings that are of width, X, and depth, X. X is uniformly distributed by 0 and 10 meters. What is the distribution of the area X^2 of the building?

$X \sim \mathcal{U}(0, 10) \Rightarrow$

$f_X(x) = \frac{1}{10}$

$F_X(x) = \int_0^x \frac{1}{10}\,dt = \frac{t}{10}\Big|_0^x = \frac{x}{10}$

$F_X(3) = \frac{3}{10}$

$F_Y(y)$ where $Y = X^2$

$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(X \le \sqrt{y}) = \frac{\sqrt{y}}{10}$

Expectation of Functions of Random Variables

Often we want to compute the expectation of a function of a random variable, instead of the random variable itself.

e.g. $E[X^2]$ instead of $E[X]$

Example: Suppose an architect is designing a community and wants to maximize the diversity in the size of his square buildings that are of width, X, and depth, X. X is uniformly distributed by 0 and 10 meters. What is the distribution of the area X^2 of the building?

So, $F_Y(y) = \frac{\sqrt{y}}{10}$ and $Y = X^2$

$$\Rightarrow \quad f_Y(y) = \frac{d}{dy}\left(\frac{\sqrt{y}}{10}\right) = \frac{d}{dy}\left(\frac{y^{1/2}}{10}\right) = \frac{\frac{1}{2}y^{-1/2}}{10}$$

$$\star A \quad \boxed{f_Y(y) = \frac{1}{20\sqrt{y}}} \quad Y \sim \frac{1}{20\sqrt{y}}$$

Expectation of Functions of Random Variables

Example: Suppose an architect is designing a community and wants to maximize the diversity in the size of his square buildings that are of width, X, and depth, X. X is uniformly distributed by 0 and 10 meters. What is the expected area?

Recall $Y = X^2$

$0 \le x \le 10$

$0 \le y \le 100$

$$E[Y]^{*} = \int_0^{100} y \cdot \frac{1}{20\sqrt{y}}\, dy$$

$$= \int_0^{100} \frac{\sqrt{y}}{20}\, dy$$

$$= \frac{y^{3/2}}{3/2} \cdot \frac{1}{20} \Big|_0^{100}$$

$$= \frac{2y^{3/2}}{3} \cdot \frac{1}{20} \Big|_0^{100} = \frac{y^{3/2}}{30} \Big|_0^{100} = \frac{100^{3/2}}{30} \boxed{= \frac{1000}{30}}$$

Expectation of Functions of Random Variables

Example: Suppose an architect is designing a community and wants to maximize the diversity in the size of his square buildings that are of width, X, and depth, X. X is uniformly distributed by 0 and 10 meters. What is the expected area?

$$E[Y] = E[X^2] = \int_0^{10} x^2 \cdot \frac{1}{10}\, dx$$

$$= \frac{X^3}{3} \cdot \frac{1}{10} \Big|_0^{10}$$

$$= \frac{X^3}{30} \Big|_0^{10}$$

$$\boxed{= \frac{1000}{30}}$$

Expectation of Functions of Random Variables

Change-of-Variables Formula: Let X be a random variable and let g: $\mathbb{R} \to \mathbb{R}$ be a function.

If X is discrete and take the values $a_1, a_2, \ldots$ then

$$E[g(x)] = \sum_i g(a_i) P(X = a_i)$$

If X is continuous, with probability density function f, then

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx$$

Expectation of Functions of Random Variables

Expectation is a Linear Function.

$$E[aX + b] = aE[X] + b$$

$$E[aX+b] = \int_{-\infty}^{\infty} (aX+b)f(x)\,dx$$

$$= \int_{-\infty}^{\infty} axf(x) + bf(x)\,dx$$

$$= \int_{-\infty}^{\infty} axf(x)\,dx \int_{-\infty}^{\infty} bf(x)\,dx$$

$$= a\int_{-\infty}^{\infty} xf(x)\,dx + b\int_{-\infty}^{\infty} f(x)\,dx$$

$$= a\,E[X] + b \cdot 1$$

$$= a\,E[X] + b$$

$f(x)$ is a pdf.

$\Rightarrow \int_{-\infty}^{\infty} f(x)\,dx = 1$

Next Time:

- Variance