9-30: Expected Value
CSCI 3022 Fall 19

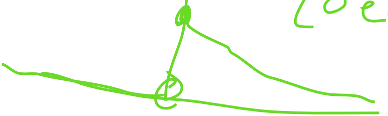
$$f(x) = \lambda e^{-\lambda x} \quad \boxed{x \geq 0}$$

Example:
Suppose a light bulb's lifetime is exponentially distributed with parameter λ.

One (often) appealing property of the exponential is it's *memoryless property*. In particular, consider the knowledge gained by knowing that the "event" has not yet occurred by time t_0. What is $P\left(X > (t_0 + t) | X > t_0\right)$?

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

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{else} \end{cases}$$



'Memoryless'

A) $P(X>(t+t_0) \mid X>t_0) = \dfrac{P(\text{both})}{P(X>t_0)} = \dfrac{P(X>(t+t_0))}{P(X>t_0)}$

$F(x) = \int_{-\infty}^{x} \overbrace{\lambda e^{-\lambda t}}^{f(x)}\, dt \underset{u=-\lambda t}{=} -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$

0 comes from support $x > 0$

$P(X \le x) = 1 - P(X > x)$

A) $= \dfrac{1 - F(t+t_0)}{1 - F(t_0)} = \dfrac{\cancel{1-(1-}e^{-\lambda(t+t_0)})}{\cancel{1-(1-}e^{-\lambda(t_0)})}$

$P(X > t) = e^{-\lambda(t+t_0) - -\lambda(t_0)} = e^{-\lambda t}$

'Memoryless'

$$P\left(X > (t_0 + t) | X > t_0\right) = \frac{P\left(X > (t_0 + t) \textit{ and } X > t_0\right)}{P\left(X > t_0\right)}$$

then use that $F(x) = 1 - e^{-\lambda x}$:

$$= \frac{1 - (1 - e^{\lambda(t_0+t)})}{1 - (1 - e^{\lambda t_0})}$$

$$= \frac{e^{\lambda(t_0+t)}}{e^{\lambda t_0}} = e^{\lambda t} = P(X > t)$$

Or we've gained no knowledge about future burnout time of the light based on the past t_0!

Announcements and Reminders

- quizlet03 for Wednesday.
- Midterm 1; Tuesday, October, 8. through W lecture
- 6:30-8:00 PM on Tuesday 8 October, this section in HUMN 1B50
  1. You are allowed to use a calculator. No smartphones or other devices that can store large amounts of data or access the internet.
  2. You are allowed one 3x5-inch notecard as a cheat sheet. You can *write* whatever you want on it and can use both sides.
  3. You do not need to bring blue books or anything like that.
  4. Do bring your Buff OneCard.
  5. Do bring multiple writing utensils.
  6. Get there early. If you arrive late, you will not receive extra time.
  7. Inform me as soon as possible about accommodations.
  8. Study from Homework, past exams, etc.

Probability Recaps

1. **Discrete:** find probabilities in the probability mass function

$$f(x) = P(X = x).$$

2. **Continuous:** find probabilities by integrating the probability density function

$$\int_a^b f(x)\,dx = P(a < X < b).$$

3. We can find cumulative probabilities or probability on ranges of outcomes in the cumulative density function

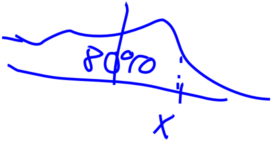
$$F(x) = P(X \le x) = \sum_{X \le x} f(x) \text{ or } \int_{-\infty}^{x} f(t)\,dt$$

Both

4. **Definition**: The median $\tilde{x}$ of a continuous distribution is the 50th percentile or $.5$ *quantile* of the distribution.
$\tilde{x}$ satisfies $F(\tilde{x}) = \boxed{.5,}$ or

any number
from [0, 1].

$$.5 = \int_{\infty}^{\tilde{x}} f(x)\,dx$$



Pops and Samples

Today marks the start of a large jump in how we approach data science problems:

1. We know about *sample statistics* like $\bar{X}$, s_X.
2. We've defined some *processes* that gives rise to distributions like the binomial, exponential, etc.
3. **Now:** we start bridging the gap! Given data and sample statistics, how do we estimate or infer properties of the underlying reality process? (parameters like p, λ).

   To do this, we need an understanding of centrality and dispersion of a process or density function might be.

Mean/Expected Value

Example:
Consider a university having 15,000 students and let X equal the number of courses for which a randomly selected student is registered.
The pdf of X is given to you as follows:

x	1	2	3	4	5	6	7
$f(x) = P(X = x)$	.01,	.03	.13	.25	.39	.17	.02

Students pay more money when enrolled in more courses, and so the university wants to know what the *average* number of courses students take per semester.

Mean/Expected Value

Definition: *Expected Value:*
For a discrete random variable X with pdf f(x), the *expected* value or *mean* value of X is denoted as E(X) and is calculated as:

$$E(X) = \sum_{x \in \Omega} x \cdot \underbrace{P(X = x)}_{f(x)}$$

outcome

prob of observing it

Compare to

$$\bar{X} = \frac{\sum x}{n}$$

count

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

outcome · "probability" of that outcome

Mean/Expected Value

Definition: *Expected Value:*
For a discrete random variable X with pdf f(x), the *expected* value or *mean* value of X is denoted as E(X) and is calculated as:

$$E[X] = \sum_{x \in \Omega} x \cdot P(X = x)$$

Mean/Expected Value

Example:, cont'd:
The pdf of X is given to you as follows:

x	1	2	3	4	5	6	7
$f(x) = P(X = x)$	.01,	.03	.13	.25	.39	.17	.02

What is $E[X]$?

$\sum$ outcomes $\cdot$ P(outcomes)

$= 1 \cdot .01 + 2 \cdot .03 + 3 \cdot .13 + 4 \cdot .25 + 5 \cdot .39$
$+ 6 \cdot .17 + 7 \cdot .02 = 4.57$

Mean/Expected Value

Example:, cont'd:
The pdf of X is given to you as follows:

x	1	2	3	4	5	6	7
$f(x) = P(X = x)$	.01,	.03	.13	.25	.39	.17	.02

What is $E[X]$?

$$E[X] = \sum_{x \in \Omega} x \cdot P(X = x) = 1 \cdot .01 + 2 \cdot .03 + 3 \cdot .13 + 4 \cdot .25 + 5 \cdot .39 + 6 \cdot .17 + 7 \cdot .02$$

$$E[X] = 4.57$$

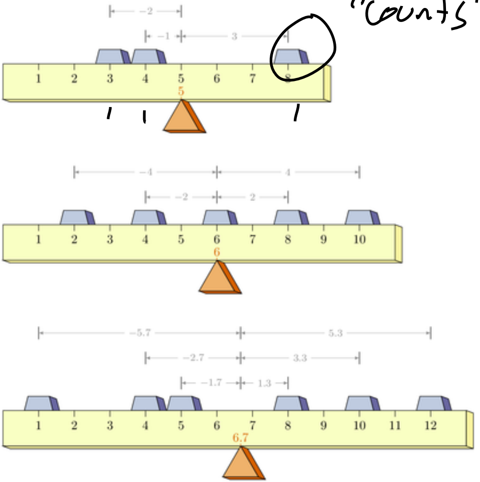
Interpreting Expected Value: Relative Frequency

One way to interpret expected value of a discrete distribution (especially on a finite support) is the sample mean if we managed to observe observations that *exactly* mirror the probability mass function.

In the preceding example, the pmf was given at 7 values of X with a precision up to 1%. In this case, if we had exactly 100 students and their proportions *observed* exactly mirrored the probabilities given in the example, the sample mean would be identical to the population mean.

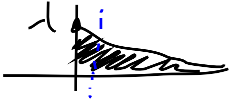
Interpreting Expected Value

further away "counts" more.



- The "center of mass" of a set of point masses
- Each mass exerts an "$r \times f$" force on the balancing point.
- Same idea holds in continuous space: we're looking for a centroid of an object.

http://www.texample.net/media/tikz/examples/TEX/balance.tex



Mean/Expected Value

Definition: *Expected Value:*
For a continuous random variable X with pdf $f(x)$, the *expected* value or *mean* value of X is denoted as $E(X)$ and is calculated as:

$$\int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

or

$$\int_{x \in \Omega} x\, f(x)\,dx$$

Mean/Expected Value

$E[\text{random vars}] \to \#$

$(\text{functions}) \to \#$

Definition: *Expected Value:*

For a continuous random variable X with pdf $f(x)$, the *expected* value or *mean* value of X is denoted as $E[X]$ and is calculated as:

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

Mean/Expected Value

$f(x) = 1/4 \cdot e^{-x/4} \quad x \geq 0$

Example:

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

$\int x e^{x} dx$

How long, on average, will the battery last? $= E[X]$

$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x \cdot 1/4 \, e^{-x/4} dx$

IBP: $\int u \, dv = uv - \int v \, du$; u = fns that change when differentiated

dv = rest.

$u = x/4 \quad v = e^{-x/4}(-4)$

$du = 1/4 \, dx \quad dv = e^{-x/4} dx$

$E[X] = -x e^{-x/4} \Big|_0^{\infty} - \underbrace{\int -e^{-x/4} dx}$

$= -x e^{-x/4} \Big|_{x=0}^{x=\infty} - 4 e^{-x/4} \Big|_{x=0}^{x=\infty} = -\,-4e^0 = \boxed{4}$

L'H & limit

Mean/Expected Value

Example:

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

How long, on average, will the battery last? Start with
$E[X] = \int_{\infty}^{\infty} x f(x)\, dx$, then use our known $f(x)$:
$E[X] = \int_0^{\infty} \lambda x e^{-\lambda x}\, dx$, now via IBP with $u = \lambda x$; $dv = e^{-\lambda x}$:
$E[X] = \lambda x (\frac{-1}{\lambda} e^{-\lambda x})|_0^{\infty} - \int_0^{\infty} \lambda (\frac{-1}{\lambda} e^{-\lambda x})\, dx$
Both xe^{-x} and $e^{-x} \to 0$ as $x \to \infty$, so we're left with:
$E[X] = \frac{-1}{\lambda} e^{-\lambda x}|_0^{\infty}$ which is just $1/\lambda$. This should come as no surprise, since we interpret λ as an average *rate* in events-per-time, but the exponential measures time-until-event, so the expected value of the exponential is the reciprocal of the rate!

Mean/Expected Value

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

How long, on average, will the battery last?

Choose for u: Log. Inverse Trig. Polynomials Exponentials Trig functions

Recall: Integration by Parts: $\int u\, dv = uv - \int v\, du$. Mental shortcuts: "integration product rule," "LIPET"

Expected Value of a Function

If a discrete r.v. X has a a density $P(X = x)$, then the expected value of any function $g(X)$ is computed as:

1. Continuous: $E[g(X)] = \int g(x) \cdot f(x)\,dx$

2. Discrete: $E[g(X)] = \sum_{x \in \Omega} g(x) \cdot \underbrace{P(X = x)}_{f(x)}$

Note that $E[g(X)]$ is computed in the same way that $E(X)$ itself is, except that $g(x)$ is substituted in place of x.

Expected Value of a Function

If a discrete r.v. X has a a density $P(X = x)$, then the expected value of any function $g(X)$ is computed as:

1. Continuous:

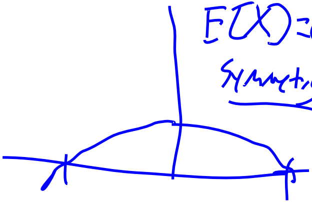
$$E[X] = \int_{\infty}^{\infty} x f(x)\ dx$$

2. Discrete:

$$E[X] = \sum_{x} x f(x)\ dx$$

Note that $E[g(X)]$ is computed in the same way that $E(X)$ itself is, except that $g(x)$ is substituted in place of x.

Expected Value of a Function



$E(X) = 0$

Symmetric

Example: A random variable X has pdf:

$$f(x) = \frac{3}{4}(1 - x^2); \quad -1 \le X \le 1$$

What is $E(X^3)$?

$$E(X^3) = \int_{-1}^{1} x^3 f(x)\,dx = \int_{-1}^{1} \frac{3}{4}x^3 - \frac{3}{4}x^5\,dx = 0$$

odd function

Symmetric region

$$= \frac{3x^4}{4\cdot 4} - \frac{3x^6}{4\cdot 6}\Big|_{-1}^{1}$$

Review: What is $F(x)$?

$$F(x) = \int_{-1}^{x} f(t)\,dt = \frac{3t}{4} - \frac{3t^3}{12}\Big|_{-1}^{x}$$

$$= \int_{-1}^{x} \frac{3}{4} - \frac{3}{4}t^2\,dt$$

Expected Value of a Function

Example: A random variable X has pdf:

$$f(x) = \frac{3}{4}(1 - x^2); \quad -1 \leq X \leq 1$$

What is $E(X^3)$?

$$E(X^3) = \int_{-1}^{1} x^3 \frac{3}{4}(1 - x^2)\, dx = \frac{3x^4}{16} - \frac{3x^6}{24}\Big|_{-1}^{1} = 0$$

Review: What is $F(x)$?

$$F(x) = \int_{-1}^{x} f(t)\, dt = \frac{3t}{4} - \frac{3t^3}{12}\Big|_{-1}^{x}$$

Expected Value of a Linear Function

If $g(X)$ is a linear function of X (i.e., $g(X) = aX + b$) then $E[g(X)]$ can be easily computed from $E(X)$.

Theorem:
Let $a, b \in \mathbb{R}$ and X be a random variable with density f. Then:

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

Proof:

Note: This works for continuous and discrete random variables.

Expected Value of a Linear Function

If $g(X)$ is a linear function of X (i.e., $g(X) = aX + b$) then $E[g(X)]$ can be easily computed from $E(X)$.

Theorem:
Let $a, b \in \mathbb{R}$ and X be a random variable with density f. Then:

$$E[g(X)] = g(E[X])$$
$$E[aX + b] = aE[X] + b$$

Proof:
$E[aX + b] = \int (ax + b)f(x)\,dx = a\int xf(x)\,dx + b\int f(x)\,dx = aE[X] + b$,
since integration is also linear!
Note: This works for continuous and discrete random variables.

Linear Expectation

Example:
Consider a university having 15,000 students and let X equal the number of courses for which a randomly selected student is registered.
The pdf of X is given to you as follows:

x	1	2	3	4	5	6	7
$f(x) = P(X = x)$	.01,	.03	.13	.25	.39	.17	.02

Earlier, we calculated that $E(X)$ was 4.57. If students pay \$500 per course plus a \$100 per-semester registration fee, what is the average amount of money the university can expect a student to pay each a semester?

$$E[\$] = E[500 \cdot \text{courses} + 100]$$
$$= 500 \cdot E[\text{courses}] + 100$$
$$= 500 \cdot 4.57 + 100$$

Linear Expectation

Example:
Consider a university having 15,000 students and let X equal the number of courses for which a randomly selected student is registered.
The pdf of X is given to you as follows:

x	1	2	3	4	5	6	7
$f(x) = P(X = x)$	.01,	.03	.13	.25	.39	.17	.02

Earlier, we calculated that $E(X)$ was 4.57. If students pay \$500 per course plus a \$100 per-semester registration fee, what is the average amount of money the university can expect a student to pay each a semester?
$Money = 500 \cdot Courses + 100 = 500X + 100 = g(X)$. Then,

$$E[g(X)] = g(E[X]) = 500 \cdot 4.57 + 100 = 2385.$$