Expectation

MATH/STAT 394: Probability I
Summer 2021 A Term

Introduction to Probability
D. Anderson, T.Seppäläinen, B. Valkó

§ 3.3

Aaron Osgood-Zimmerman

Department of Statistics

Logistics

- take home midterm, available at 10:45 am today through 11:59am tomorrow
  - open book
  - no time limit within the 25 hours
  - submit solutions to gradescope in pdf form
  - latex template available
  - work by yourself
- homework 3 deadline extended to july 10th at 11:59pm
- homework 4 will be available tonight or tomorrow morning, due july 16th
- I re-opened and extended the time to fill out the mid-term feedback to allow more responses. please consider filling it out

Practice solution

Practice

John has an insurance policy on his car with a 200$ deductible, meaning that if an accident occur, he would pay the cost of the repair up to 200$ with the insurance policy paying the rest. So if he has an accident worth 123$ he would pay 123$ but if the accident is worth 345$ he would only pay 200$.

Assume that the cost of an accident is uniformly distributed over $[50, 1000]$. Denote X the amount that John pays.

- What is the c.d.f. of X?
- Is X continuous, discrete or neither discrete nor continuous?
- What is the prob. that $X = 200$?

Solution

- Since no accident is worth less than 50$, we have that for $t \leq 50$, then $F(t) = \mathbb{P}(X \leq t) = 0$
- If $t \geq 200$, then since John pays at most 200, we have that $\mathbb{P}(X \leq t) = \mathbb{P}(X \leq 200) = 1$.
- Now denote $Y \sim \text{Unif}([50, 1000])$ the cost of the accident, recall that
  - p.d.f. $f(t) = \frac{1}{1000-50}$ for $t \in [50, 1000]$ and 0 o.w.
  - c.d.f. $F(t) = \frac{t-50}{1000-50}$ for $t \in [50, 1000]$
- If $t \in [50, 200]$, then John pays the amount he gave, so $\{X \leq t\} = \{Y \leq t\}$ and we know the c.d.f. of a uniform r.v. m namely

$$F(t) = \mathbb{P}(X \leq t) = \mathbb{P}(Y \leq t) = \frac{t-50}{950}$$

Practice solution

Practice

John has an insurance policy on his car with a 200$ deductible, meaning that if an accident occur, he would pay the cost of the repair up to 200$ with the insurance policy paying the rest. So if he has an accident worth 123$ he would pay 123$ but if the accident is worth 345$ he would only pay 200$.

Assume that the cost of an accident is uniformly distributed over $[50, 1000]$. Denote X the amount that John pays.

- What is the c.d.f. of X?
- Is X continuous, discrete or neither discrete nor continuous?
- What is the prob. that $X = 200$?

Solution (Continued)

- In summary

$$F(t) = \begin{cases} 0 & \text{if } t < 50 \\ \frac{t-50}{950} & \text{if } 50 \leq t < 200 \\ 1 & \text{if } 200 \leq t \end{cases}$$

- It is neither discrete (not piece-wise constant), nor continuous (since it it is discontinuous at 200)
- We have that

$$\mathbb{P}(X = 200) = F(200) - F(200-) = 1 - \frac{150}{950} = \frac{16}{19} \approx 0.84$$

Outline

Summary

c.d.f. defined for any r.v. X

$$F(t) = \mathbb{P}(X \leq t)$$

From p.m.f./p.d.f. to c.d.f.

$$F(t) = \sum_{\substack{k \in \mathcal{X} \\ k \leq t}} p(k) \qquad \text{(for discrete r.v.)}$$

$$F(t) = \int_{-\infty}^{t} f(x) dx \qquad \text{(for continuous r.v.)}$$

From c.d.f. to p.m.f./p.d.f.

$$p(k) = F(k+) - F(k-) \qquad \text{(for discrete r.v.)}$$

$$f(x) = F'(x) \qquad \text{(for continuous r.v.)}$$

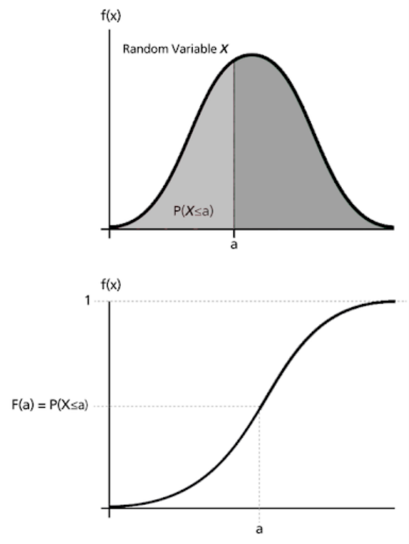
pdf and cdf graphically

c.d.f. defined for any r.v. X

$$F(t) = \mathbb{P}(X \leq t)$$

From p.m.f./p.d.f. to c.d.f.

$$F(t) = \int_{-\infty}^{t} f(x)dx \quad \text{(for continuous r.v.)}$$



Summary

Charac. of random variables by the c.d.f.

- $F(t) = \mathbb{P}(X \leq t)$ is
  - non-decreasing
  - $F(-\infty) = 0$, $F(+\infty) = 1$
  - right continuous
  - $F(t-) = \mathbb{P}(X < t)$
- If F is piece-wise constant
  $\rightarrow$ it is the c.d.f. of a **discrete r.v.**
- If F is continuous
  $\rightarrow$ it is the c.d.f. of a **continuous r.v.**
- If F is discontinuous and not piece-wise constant
  $\rightarrow$ neither discrete nor continuous r.v.
  Yet, we can compute prob. using the c.d.f.

Why did we introduce the c.d.f.?

Theoretical reason

- We only need $\mathbb{P}(X \leq t)$ for any t to compute any prob. measure
- Therefore the c.d.f. is a priori sufficient for our purposes

Practical reason

- The c.d.f. is a prob. so we can use classical rules of prob. to manipulate it
- On the other hand the p.d.f. is just a function and it may be sometimes not practical

Why did we introduce the c.d.f.?

Exercise

Let $X \sim \text{Exp}(\lambda)$, $Y \sim \text{Exp}(\mu)$ be independent.

What is the p.d.f. of $M = \min(X, Y)$?

Recall that $F_X(t) = 1 - e^{-\lambda t}$.

Solution

- We have that the event $\{\min\{X, Y\} > t\}$ is equivalent to the event $\{X > t\} \cap \{Y > t\}$.
- Therefore

$$\begin{aligned} 1 - F_M(t) = \mathbb{P}(\min(X, Y) > t) &= \mathbb{P}(X > t, Y > t) \\ &= \mathbb{P}(X > t)\mathbb{P}(Y > t) \qquad \text{(by independence)} \\ &= e^{-\lambda t} e^{-\mu t} = e^{-(\lambda+\mu)t} \end{aligned}$$

- So the p.d.f of M is given as

$$f_M(x) = F_M'(x) = (1 - e^{-(\lambda+\mu)x})' = (\lambda + \mu)e^{-(\lambda+\mu)x}$$

- In fact we recognize that $M \sim \text{Exp}(\lambda + \mu)$.

Random generation

Motivation

- We defined some r.v.theoretically but can we have concrete realizations of these r.v.
- Namely how could we **generate some randomness** ?

Modern computers give you access to a **Random Number Generator**

"Generates a sequence of numbers
that cannot be reasonably predicted better than by a random chance."

It typically means that have access to uniform distribution U on $[0, 1]$

Random generator

Example (Sampling from a non-uniform dist.)
Assume that we have access to a c.d.f. F_X of a r.v. X (e.g. an exponential r.v.) with F *strictly increasing*

How can we have access to some random realizations of X, i.e., we want the realizations to be such that they follow the distribution of X?

Solution

- Define $T = F_X^{-1}$ (possible if F strictly increasing)
- Generate a random point $U \sim \mathrm{Unif}([0,1])$ and define $Y = T(U)$
- Note that for any $a \in [0,1]$,

$$\mathbb{P}(U \leq a) = \int_0^a dx = a$$

- Then

$$\mathbb{P}(Y \leq t) = \mathbb{P}(U \leq F_X(t)) = F_X(t) = \mathbb{P}(X \leq t)$$

- So the number we generated has the same distribution as the one of X!

Outline

Expectation

Motivation

- Given a r.v. we have numerous tools to compute prob.
- Cool, but beyond computing prob., what are the key properties we would like to know about a r.v.?
- Typically if you flip a coin n times, you would like to know what is the average number of tails you should get
- In prob., this average number is called an **expectation** and it is a measure of central tendency

Intuition

Example
At a casino,

- you lose 1$ 90% of the time
- you gain 10$ 9% of the time
- you gain 100$ 1% of the time

What is your expected net gain?

Solution

- First understand that the average is a **number** not a prob.
- Then

$$\text{expected net gain} = \underbrace{(-1)}_{\text{net gain}} \cdot \underbrace{\frac{90}{100}}_{\text{frequency}} + 10 \cdot \frac{9}{100} + 100 \cdot \frac{1}{100} = 1$$

Outline

Expectation of discrete r.v.

Definition

Let X be a discrete r.v. with p.m.f. p, its ***expectation*** *is defined as*

$$\mathbb{E}[X] = \sum_{k \in \mathcal{X}} k \mathbb{P}(X = k) = \sum_{k \in \mathcal{X}} k p(k)$$

where $\mathcal{X}$ is the set of values that X can take with non zero prob.

The expectation a.k.a. mean is often denoted μ_X.

Note:

- The expectation may be finite, infinite or not defined (see examples later)

Expectation of discrete r.v.

Exercise

Flip a biased coin (with $\mathbb{P}(T) = p$) n times, what is the expected number of tails you get? In other words, compute $\mathbb{E}[X]$ for $X \sim \text{Bin}(n, p)$

Hint: recall $(x + y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \ldots + \binom{n}{n}y^n = \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k$

Solution

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k\mathbb{P}(X = k) = \sum_{k=1}^{n} k\mathbb{P}(X = k) \\
&= \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}(1-p)^{n-k} \\
&= np\sum_{j=0}^{n-1}\frac{(n-1)!}{j!(n-1-j)!}p^j(1-p)^{n-1-j}, \quad \text{(j=k-1)}, \\
&= np\sum_{j=0}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-1-j} = np\,[p + (1-p)]^{n-1}, \quad \text{(binomial theorem)} \\
&= np1^{n-1} = np.
\end{aligned}
$$

Expectation of a discrete r.v.

Example
Flip a biased coin ($\mathbb{P}(X = T) = p$) until you get a tail. What is the expected number of times you need to flip the coin?

In other words, compute $\mathbb{E}[X]$ for $X \sim \text{Geom}(p)$

Solution

- A power series can be differentiated inside its radius of convergence. Namely we have for $|t| < 1$,
$$\sum_{k=0}^{+\infty} t^k = \frac{1}{1-t}$$
so
$$\sum_{k=1}^{+\infty} kt^{k-1} = \sum_{k=0}^{+\infty} kt^{k-1} = \frac{d}{dt}\left(\sum_{k=0}^{+\infty} t^k\right) = \frac{d}{dt}\frac{1}{1-t} = \frac{1}{(1-t)^2}$$
- Therefore
$$\mathbb{E}[X] = \sum_{k=1}^{+\infty} k(1-p)^{k-1}p = p\frac{1}{(1-(1-p))^2} = \frac{1}{p}$$

Expectation of a discrete r.v.

Exercise

Let $X \sim \text{Ber}(p)$, what is $\mathbb{E}[X]$?

Solution

- $\mathbb{E}[X] = 1 \cdot p + 0 \cdot (1 - p) = p$

Link between expectation and prob.

- For an event $A \subseteq \Omega$ recall the definition of the indicator function of A,

$$1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

- $1_A \sim \text{Ber}(\mathbb{P}(A))$ (since $\mathbb{P}(1_A = 1) = \mathbb{P}(\omega \in A)$)
- Therefore

$$\boxed{\mathbb{E}[1_A] = \mathbb{P}(A)}$$

Outline

Expectation of continuous r.v.

Definition

*Let X be a continuous r.v. with p.d.f. f. Then its **expectation** is defined as*

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

The expectation a.k.a. mean is often denoted μ_X.

Notes

- As before, it may be finite/infinite or not defined
- Note that compared to the discrete case we simply moved from a sum to an integral

Interpretation

- The expectation of X can be seen as "the center of mass" of the dist.
- The center of mass of an object is indeed defined by integrating the density times the position

Expectation of a continuous r.v.

Exercise

Draw a point uniformly at random on $[a, b]$ with $a < b$, what is the expected value of this point?

In other words, compute $\mathbb{E}[X]$, for $X \sim \text{Unif}([a, b])$

Solution

- $$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \left[\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$
- Note that is make sense: it is just the middle of the interval.
- If $[a, b]$ was a rope with uniform density, its center of mass is the middle of the rope.

Practice next lecture

Practice

The time (in min.) to wait for my bus can be modeled as an exponential r.v.

I know that the average time I need to wait for the bus is 10 min.

What is the prob. that I would wait more than 15min?

Hint:

1. Compute $\mathbb{E}[X]$ for $X \sim \mathrm{Exp}(\lambda)$.
2. Deduce how you could know λ if you only knew the expected time.

Note: We could do a similar exercise with a geometric r.v.:

if we know the mean then we know the param. of the dist.