1.) definition of expected value

a.)

$\mathbb{E}[X] = \sum_{k=1}^{\infty} x_k P(X = x_k) \rightarrow$ discrete

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

In the continuous case, expectations are not always defined because the $\int_{-\infty}^{\infty} x f(x) dx$ may go to infinity

For example, the cauchy distribution has a pdf of $\frac{1}{\pi} \cdot \frac{1}{1+x^2}$ $\quad -\infty < x < \infty$

Given our continuous def'n of expected value, $\mathbb{E}[X] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x}{1+x^2} dx = \frac{1}{\pi} \left(\frac{\ln(x^2+1)}{2} \right) \Big|_{-\infty}^{\infty}$ = undefined

b.)

$\mathbb{E}[X] = \sum_{k=1}^{\infty} x_k P(X = x_k) = \sum_{k=0}^{\infty} \sum_{j=0}^{k} p(X = k) \longrightarrow a + 2b + 3c + 4d$

$$\begin{aligned} a &+ b + c + d \ldots \\ &+ b + c + d \ldots \\ &\quad\;\; + c + d \ldots \\ &\qquad\quad\; + d \ldots \end{aligned}$$

$\sum_{k=0}^{\infty} P(X > k) = \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} p(X = k) = \sum_{k=0}^{\infty} \sum_{j=0}^{k} p(X = k) = \mathbb{E}[X]$

every prob $\to j$

Geometric variable

$p(X = x) = (1-p)^{x-1} \cdot p$

$p(X > x) = (1-p)^x$

this is a geometric series

$\mathbb{E}[X] = \sum_{k=0}^{\infty} p(X > k) = \sum_{k=1}^{\infty} (1-p)^k$ so $\sum_{k=1}^{\infty} (1-p)^k = \frac{1}{1-(1-p)} = \frac{1}{p}$

a) Give the definition of the expected value for discrete and for continuous random variables and explain why expectations are not always defined. Use the $t_{(1)}$ (Cauchy) as an example, and explain why we can't just say the mean is 0 due to symmetry.

b)* Show that, for a discrete random variable X that takes only non-negative integers, $E(X) = \sum_{k=0}^{\infty} P(X > k)$. Hint: re-express k as a sum from $j = 1$ to k of 1, then reverse the order of summation. Use this to find the expected value of a Geometric(θ) random variable that counts the trials until the first success (compare to using the definition to find $E(X)$).

c) State the results for the linearity of expectation. Review how Blitzstein argues, without using joint distributions, that $E(X+Y) = E(X) + E(Y)$, even if X and Y are not independent. Give examples using extremely correlated variables (e.g., $X_1 = X_2$ and $X_1 = -X_2$), Show that if $E(X_i) = \mu$ for $i = 1, \ldots, n$, then $E(\bar{X}) = \mu$.

d) Prove Markov's and Chebyshev's inequalities, and show how these imply the weak law of large numbers. Use the fact that, for $X_1, \ldots, X_n$ iid, $\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$.

c.)

Results for linearity of expectation:

1.) $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

2.) $\mathbb{E}[cX] = c\mathbb{E}[X]$

proving 2:

$\mathbb{E}[cX] = \sum_{i=0}^{\infty} c x_i \cdot p(X = x_i) = c \sum_{i=0}^{\infty} x_i \cdot p(X = x_i) = c\mathbb{E}[X]$ ✓

proving 1:

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \leftarrow$ obvious when X and Y are independent

What if dependant?

Suppose $X = Y$ $\leftarrow$ most extreme case of dependence

then $\mathbb{E}[X+Y] = \mathbb{E}[2X] = 2\mathbb{E}[X]$

and $\mathbb{E}[X] + \mathbb{E}[Y] = 2\mathbb{E}[X]$

So $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

What if $X = -Y$:

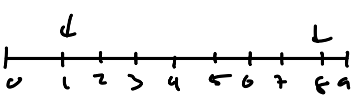
then $\mathbb{E}[X-Y] = \mathbb{E}[0] = 0$

and $\mathbb{E}[X] + \mathbb{E}[-Y] = \mathbb{E}[X] - \mathbb{E}[Y] = \mathbb{E}[X] - \mathbb{E}[X] = 0$

Show that if $\mathbb{E}[X_i] = \mu$ for $i = 1, \ldots, n$, then $\mathbb{E}[\bar{X}] = \mu$

$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \Rightarrow \mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right] = \frac{1}{n} \cdot \mathbb{E}\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[x_i] = \frac{1}{n} \sum_{i=1}^{n} \mu = \mu$

d.)



Markov's

$p(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$ for non-neg. X

proof!

$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{\infty} x f(x) dx$ because x is non-negative

$\int_0^{\infty} x f(x) dx = \int_0^a x f(x) dx + \int_a^{\infty} x f(x) dx \geq \int_a^{\infty} x f(x) dx \geq \int_a^{\infty} a f(x) dx \geq a \int_a^{\infty} f(x) dx = a \cdot p(X \geq a)$

always lower bound

So $\mathbb{E}[X] > a \cdot p(X \geq a) \rightarrow p(X \geq a) < \frac{\mathbb{E}[X]}{a}$

Chebyshev

$p(|x-\mu| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}$

proof:

we know $p(|x-\mu| > \varepsilon) = p((x-\mu)^2 > \varepsilon^2) \leq \frac{\mathbb{E}[(x-\mu)^2]}{\varepsilon^2} = \frac{\text{Var}(x)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$

Weak law of large numbers

$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\text{var}(\bar{x}_n) = \frac{\sigma^2}{n}$

law states that $P(|\bar{x}_n - \mu| \geq \varepsilon) \to 0$ as $n \to \infty$ for any $\varepsilon > 0$

proof using Chebyshev

$p(|\bar{x}_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\bar{x}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$

and $\lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$ so $p(|\bar{x}_n - \mu| \geq \varepsilon) \xrightarrow{n\to\infty} 0$