MATH 115
LECTURE 17

: Expected Value, Variance / Continuous Random Variables.

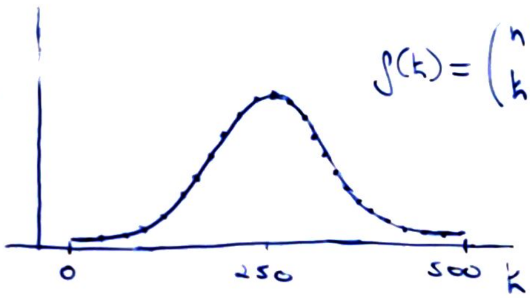
The previous lecture we introduce random variables as a way to study events. Associated to every random variable we have its probability distribution, which summarizes how likely each possible outcome is.

Ex: X = number of heads when tossing a coin 500 times.

so X is a binomial random variable with $n=500$, $p=\frac{1}{2}$.

Its distribution looks like

$P(X=k) = f(k)$



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

• Moreover, we learn how to compute the expected value of a discrete random variable whose distribution is $f(x)$:

$$E(X) = \sum_k x_k f(x_k).$$

→ In the particular case of a binomial random variable,

$E(X) = np$ (expected number of successes in n trials, when p = probability of success.)

• Our intuition about the meaning of $E(X)$ is indeed a theorem:

The Law of Large Numbers:

Consider an experiment for which we define a random variable X.
(for example: toss a coin 10 times, X = number of heads).

Now, we repeat this experiment many times, and for each time we write down the value of X.

So for example, 1st experiment $\to X_1 = 3$ (3 heads out of 10)
2nd " $\to X_2 = 6$ (6 heads out of 10)
3rd " $\to X_3 = 4$ (4 " " " ")
$\vdots$

If we do it n times, we can define a new random variable, the <u>sample mean of X</u>, $\bar{X}_n$:

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}.$$

Then, if we denote $\mu = E(X)$ (expected value of X, also called mean of X), it holds that

$$\lim_{n\to\infty} P(\mu - \varepsilon \le \bar{X}_n \le \mu + \varepsilon) = 1 \quad \text{for all } \varepsilon > 0.$$

(that is, "$\bar{X}_n \to \mu$ as $n \to \infty$")

- <u>Variance and standard deviation</u>

While $E(x)$ measures the mean value of x, the variance $Var(X)$ and the standard deviation $\sigma(x)$ measures how much x deviates from its mean $E(x)$. In other words, how likely it is that the values of x are actually near $E(x)$.

- <u>Def</u>: Let x be a discrete random variable with distribution $f(x)$ and mean $\mu = E(x)$. Then, we define

$$Var(X) = \sum_k (x_k - \mu)^2 f(x_k) = E((X-\mu)^2)$$

$$\sigma(X) = \sqrt{Var(x)}$$

<u>Remark</u>: You can show that (check it)

$$\boxed{Var(x) = E(x^2) - \mu^2}$$

- Binomial random variable: $E(x) = np$

$Var(x) = npq$, $\sigma(x) = \sqrt{npq}$.

• Example: Find the mean and variance of X.

x	1	3	4	5
$f(x)$	0.4	0.1	0.2	0.3

Sol: $\mu = E(x) = \sum_k x_k f(x_k) = 1(0.4) + 3(0.1) + 4(0.2) + 5(0.3) = 3.$

$Var(X) = E(x^2) - \mu^2 = \sum_k x_k^2 f(x_k) - \mu^2 =$

$= 1^2(0.4) + 3^2(0.1) + 4^2 \cdot (0.2) + 5^2 \cdot (0.3) - 3^2 = 12 - 9 = 3.$

• Exercise: A coin is weighted so that $P(H) = p$ and hence $P(T) = q = 1-p$. The coin is tossed until a heads appears.

→ Show that the expected number of tosses is $\frac{1}{p}$.

→ Show that the standard deviation in the number of tosses is

Sol:

Sample space: $S = \{H, TH, TTH, \ldots, T^k H, \ldots, T^\infty\}$

Random variable X = number of tosses, $R(X) = \{1, 2, 3, \ldots\}$

Probability distribution: $f(k) = P(X=k) = q^{k-1} p$

Thus,

$$E(X) = \sum_{k=1}^{\infty} k\, f(k) = \sum_{k=1}^{\infty} k\, q^{k-1} p = p \underbrace{\sum_{k=1}^{\infty} k\, q^{k-1}}_{?}.$$

Recall that for any $|q|<1$ we have that

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

$$S_n = 1 + q + q^2 + \dots + q^n$$

$$qS_n = q + q^2 + q^3 + \dots + q^n + q^{n+1},$$

$$\hookrightarrow S_n(1-q) = 1 - q^{n+1} \Rightarrow S_n = \frac{1-q^{n+1}}{1-q} \Rightarrow$$

$$\Rightarrow \lim_{n\to\infty} S_n = \frac{1}{1-q}$$

Therefore,

$$\frac{d}{dq} \sum_{k=0}^{\infty} q^k = \sum_{k=0}^{\infty} k\, q^{k-1} = \sum_{k=1}^{\infty} k\, q^{k-1} = \frac{d}{dq}\left(\frac{1}{1-q}\right) = \frac{+1}{(1-q)^2}.$$

and thus we conclude that

$$E(X) = p\,\frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}.$$

Now,

$$Var(X) = E(X^2) - \mu^2 = \sum_{k=1}^{\infty} k^2 f(k) - \frac{1}{p^2} = \underbrace{\sum_{k=1}^{\infty} k^2 q^{k-1} p}_{?} - \frac{1}{p^2}.$$

We use a similar trick:

$$\frac{d}{dq}\sum_{k=1}^{\infty} k\, q^{k} = \sum_{k=1}^{\infty} k^2 q^{k-1}$$ and we know from the previous part that

$$\sum_{k=1}^{\infty} k\, q^{k-1} = \frac{1}{(1-q)^2}, \text{ so } \sum_{k=1}^{\infty} k\, q^{k} = \frac{q}{(1-q)^2}, \text{ and thus}$$

$$\sum_{k=1}^{\infty} k^2 q^{k-1} = \frac{d}{dq}\left(\frac{q}{(1-q)^2}\right) = \frac{1}{(1-q)^2} + \frac{2q}{(1-q)^3} = \frac{1-q+2q}{(1-q)^3} = \frac{1+q}{(1-q)^3}.$$

In conclusion,

$$\mathrm{Var}(X) = p\cdot\frac{1+q}{p^3} - \frac{1}{p^2} = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2},$$

$$\sigma(X) = \frac{\sqrt{q}}{p}.$$

2.1 Continuous random variable

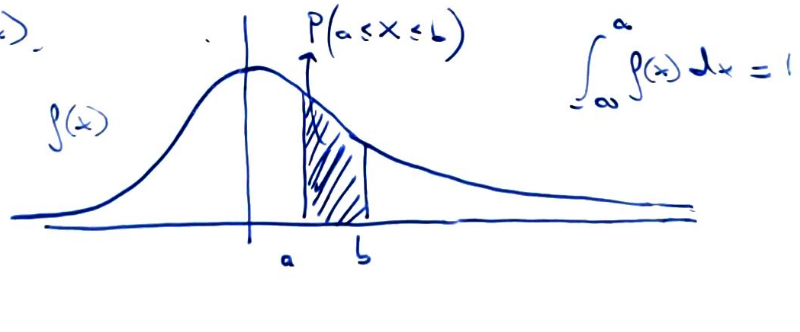
X is a <u>continuous</u> random variable if its range $R(X)$ is a continuum of numbers such as $\mathbb{R}$ or an interval.

Say that $R(X) = [a,b] \subseteq \mathbb{R}$. Then, we associate to X a <u>density function</u> $f(x)$ such that

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

The function $f(x)$ has to satisfy that
1) $f(x) \ge 0$,
2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

<u>Remark:</u> A continuous random variable is defined through its density function, and we compute the probabilities as <u>areas</u> under the graph of $f(x)$.



Notice the analogy with discrete random variables:

Here, it is meaningless to talk about the probability of a

particular value of X, since $P(X = x_0) = 0$ for all x_0
(there are "∞ possible values for X").

However, we can consider the probability of being near x_0:

$P(x_0 - \frac{1}{2}\Delta x \leq X \leq x_0 + \frac{1}{2}\Delta x) \approx f(x_0)\Delta x$.

→ In the discrete case:

$$P(a \leq X \leq b) = \sum_k P(X = x_k) = \sum_k f(x_k), \quad k = 1, 2, \dots, n$$
$$x_1 = a, \; x_n = b.$$

→ Continuous case:

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx \approx \sum_k f(x_k)\Delta x_k \approx \sum_k P\left(x_k - \frac{\Delta x_k}{2} \leq X \leq x_k + \frac{\Delta x_k}{2}\right).$$

• <u>Def</u>: Cumulative distribution of X

Given X continuous random variable with density function, we define the cumulative distribution of X as

$$F(x) = \int_a^x f(s)\,ds = P(X \leq x), \quad \text{where } R(X) = [a, b].$$
(might be $\pm\infty$)

<u>Remark</u>: 1) $P(a \leq X \leq b) = F(b) - F(a)$.

2) $F(x_1) \leq F(x_2)$ for all $x_1 \leq x_2$.

3) $\lim_{x \to a} F(x) = 0, \; \lim_{x \to b} F(x) = 1$.

• <u>Exercise</u>: Let $f(x) = \frac{1}{x^2}$ on $x > 1$.

a) Show that f is a density function.

b) Find the cumulative distribution of f.

c) Find $P(3 \leq X \leq 4)$.

<u>Sol</u>:

a) $f(x) \geq 0$, $\int_1^\infty \frac{1}{x^2}\,dx = \left.\frac{-1}{x}\right]_1^\infty = 1$.

b) $F(x) = \int_1^x f(s)\,ds = \int_1^x \frac{1}{s^2}\,ds = \left.\frac{-1}{s}\right]_1^x = 1 - \frac{1}{x}$.

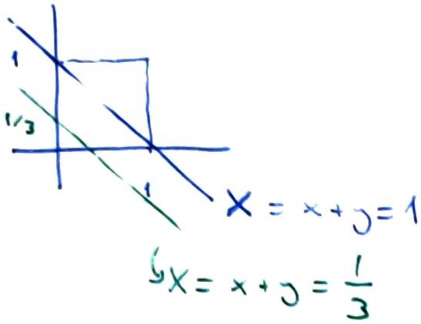
c) $P(3 \leq X \leq 4) = F(4) - F(3) = \cancel{1} - \frac{1}{4} - \cancel{1} + \frac{1}{3} = \frac{1}{3} - \frac{1}{4} = \frac{4-3}{12} = \frac{1}{12}$.

• <u>Exercise</u>: A point is chosen at random from the unit square $[0,1] \times [0,1]$. Let X be the sum of its two coordinates. Find the density function and the cumulative distribution of X. Find $P\left(\frac{1}{2} < X < 1\right)$.

<u>Sol</u>: The sample space is $S = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}$.

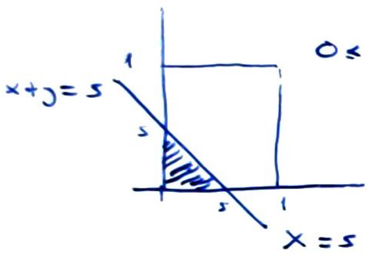
$$\left.\begin{aligned} X &: S \to [0,2] \\ (x,y) &\mapsto X((x,y)) = x+y \end{aligned}\right\}$$

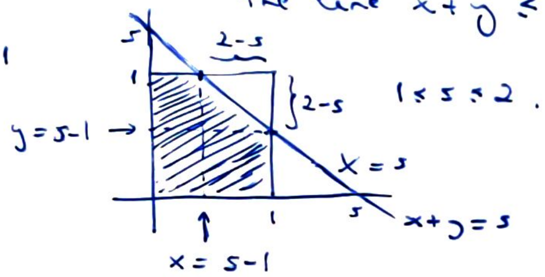
Notice that X is constant along the lines $x+y=c$: $(X=c)$



Therefore,

$F(s) = P(X \leq s) =$ "area in the square $[0,1]\times[0,1]$ under the line $x+y \leq s$".





we can see that

- For $0 \leq s \leq 1$, $F(s) = \frac{s^2}{2}$,
- For $1 \leq s \leq 2$, $F(s) = \frac{1}{2} + \frac{1}{2} - \frac{(2-s)^2}{2} = 1 - \frac{(2-s)^2}{2}$,

that is

$$F(s) = \begin{cases} s^2/2 & 0 \leq s \leq 1 \\ 1 - \frac{(2-s)^2}{2} & 1 \leq s \leq 2 \end{cases}$$

(Check: $F(2) = 1$, $F(0) = 0$).

Then, $f(s) = F'(s)$, so

$$f(s) = \begin{cases} s & 0 \leq s \leq 1, \\ 2-s & 1 \leq s \leq 2, \end{cases}$$

and
$$P\left(\tfrac{1}{2} < X < 1\right) = F(1) - F\left(\tfrac{1}{2}\right) = \frac{1}{2} - \frac{1}{8} = \frac{3}{8}.$$

- We can define $E(X)$, $Var(X)$ and $\sigma(X)$ for continuous random variables analogously to the discrete case:

$$E(X) = \int_a^b x\, f(x)\, dx, \quad Var(X) = E((X-\mu)^2) = E(X^2) - \mu^2 = \int_a^b x^2 f(x)\, dx - \mu^2,$$

where $R(X) = [a,b]$ (a, b can be $\pm\infty$),
$\mu = E(X)$.

- Example. Find μ and σ for X in the previous exercise.

Sol:
$$\mu = E(X) = \int_0^2 s\, f(s)\, ds = \int_0^1 s^2\, ds + \int_1^2 s(2-s)\, ds = 1. \quad \left(\begin{array}{c}\text{the result is} \\ \text{intuitive}\end{array}\right).$$

$$\sigma^2 = Var(X) = \int_0^2 s^2 f(s)\, ds - 1^2 = \int_0^1 s^3\, ds + \int_1^2 s^2(2-s)\, ds - 1 = \frac{7}{6} - 1 = \frac{1}{6}.$$