10/13

Last time: Expectation $E(X)$

Discrete

$$\sum_i x_i \, p(x_i)$$

Continuous.

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

Today:

Theorem: Suppose that $Y = g(X)$.

g: "Nice" function.

(a) If X is discrete with p.m.f. $p(x)$

then, $E(Y) = \sum_x g(x)\, p(x)$ $\quad \boxed{E(X) = \sum_x x\, p(x)}$

(provided $\sum_x |g(x)|\, p(x) < \infty$)

(b) If X is continuous with pdf $f(x)$

then $E(Y) = \int_{-\infty}^{\infty} g(x) f(x)\, dx$ $\quad \boxed{E(X) = \int_{-\infty}^{\infty} x f(x)\, dx}$

(provided $\int_{-\infty}^{\infty} |g(x)|\, f(x)\, dx < \infty$)

Example: X: the magnitude of the velocity of gas molecule

Then it is known (Maxwell's distribution) that

$$f_X(x) = \begin{cases} \frac{\sqrt{2/\pi}}{\sigma^3} \cdot x^2 e^{-\frac{x^2}{2\sigma^2}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

pdf

(~~Average~~ Kinetic energy =) $Y = \frac{1}{2} m X^2$

New random variable

Question: What is the average kinetic energy? $E(Y)$

Solution:

$$E(Y) = \int_{-\infty}^{\infty} \underbrace{\left(\frac{1}{2} m x^2\right)}_{y} f_X(x)\, dx$$

$$= \int_{0}^{\infty} \frac{1}{2} m x^2 \cdot \frac{\sqrt{2/\pi}}{\sigma^3} \cdot x^2 e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= \frac{\sqrt{2/\pi}}{2\sigma^3} m \int_0^{\infty} \boxed{x^4 \cdot e^{-\frac{x^2}{2\sigma^2}}} dx$$

even

$$= \frac{\sqrt{2/\pi}}{2\sigma^3} m \cdot \int_0^{\infty} x^4 e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= \frac{\sqrt{2/\pi}}{2\sigma^3}\, m \int_0^\infty x^4 e^{-x^2/2\sigma^2}\, dx$$

$$= \frac{\sqrt{2/\pi}}{2\sigma^3}\, m \int_0^\infty 4\sigma^4 u^2\, e^{-u}\, \frac{\sigma}{\sqrt{2u}}\, du$$

$$u = \frac{x^2}{2\sigma^2}$$

$$du = \frac{\cancel{2} x\, dx}{\cancel{2}\sigma^2}$$

$$dx = \frac{\sigma^2}{x}\, du$$

$$= \frac{\sigma^{\cancel{2}}}{\sqrt{2}\,\cancel{\sigma}\sqrt{u}}\, du$$

$$\Gamma(x) = \int_0^\infty e^{-u} u^{x-1}\, du$$

$$x \neq -1, -2, -3, \ldots$$

$$= \frac{\sqrt{\cancel{2}/\pi}}{2}\, \sigma^2 m\, \frac{4}{\cancel{\sqrt{2}}} \int_0^\infty u^{3/2} e^{-u}\, du$$

$$= \frac{4 m \sigma^2}{2\sqrt{\pi}} \cdot \int_0^\infty u^{\frac{5}{2}-1} e^{-u}\, du$$

$$= \frac{4 m \sigma^2}{2\sqrt{\pi}} \cdot \Gamma\left(\frac{5}{2}\right)$$

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right)$$

$$= \frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$= \frac{3}{4}\cdot\sqrt{\pi}$$

$$= \frac{\cancel{4} m \sigma^2}{2\cancel{\sqrt{\pi}}} \cdot \frac{3}{\cancel{4}}\cancel{\sqrt{\pi}}$$

$$= \boxed{\frac{3}{2} m \sigma^2}$$

//

Theorem: Suppose that $X_1, X_2, \ldots, X_n$ are jointly distributed random variables and

$$\boxed{Y = g(X_1, X_2, \ldots, X_n)}$$

(a) If the X_i are discrete with joint p.m.f. $p(x_1, x_2, \ldots, x_n)$, Then

$$\boxed{E(Y) = \sum_{x_1, x_2, \ldots, x_n} g(x_1, x_2, \ldots, x_n) \cdot p(x_1, x_2, \ldots, x_n)}$$

(b) If X_i are continuous with joint density function $f(x_1, x_2, \ldots, x_n)$, then

$$\boxed{E(Y) = \underbrace{\int\int \cdots \int}_{n\text{-times}} g(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n) \cdot dx_1 \cdot dx_2 \ldots dx_n.}$$

Corollary: If X and Y are independent random variables and g and h are fixed functions, then

$$\boxed{E[g(X)\,h(Y)] = E[g(X)] \cdot E[h(Y)]}$$

Theorem: If $X_1, X_2, \ldots, X_n$ are jointly distributed random variables with expectations $E(X_i)$ and Y is given by

$$Y = a + \sum_{i=1}^{n} b_i X_i$$

then

$$E(Y) = a + \sum_{i=1}^{n} b_i E(X_i).$$

where $a, b_1, b_2, \ldots, b_n$ are constants.

Proof: (Continuous case)

$n = 2$

$$Y = a + b_1 X_1 + b_2 X_2 = g(X_1, X_2)$$

So $g(x_1, x_2) = a + b_1 x_1 + b_2 x_2$

$$E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \underbrace{(a + b_1 x_1 + b_2 x_2)}_{g(x_1, x_2)} f(x_1, x_2)\, dx_1\, dx_2$$

$$= a \boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2}^{1} + b_1 \int_{x_2=-\infty}^{\infty}\int_{x_1=-\infty}^{\infty} x_1 f(x_1, x_2)\, dx_1\, dx_2$$

$$+ b_2 \int_{x_1=-\infty}^{\infty}\int_{x_2=-\infty}^{\infty} x_2 f(x_1, x_2)\, dx_1\, dx_2$$

$$= a + b_1 \int_{x_1=-\infty}^{\infty} x_1 \left(\int_{x_2=-\infty}^{\infty} f(x_1, x_2)\, dx_2 \right) dx_1 + b_2 \int_{x_2=-\infty}^{\infty} x_2 \left(\int_{x_1=-\infty}^{\infty} f(x_1, x_2)\, dx_1 \right) dx_2$$

$$= a + b_1 \int_{x_1=-\infty}^{\infty} x_1 f_{X_1}(x_1)\, dx_1 + b_2 \int_{x_2=-\infty}^{\infty} x_2 f_{X_2}(x_2)\, dx_2$$

pdf of X_1 (marginal pdf)

$$= a + b_1 E(X_1) + b_2 E(X_2)$$

$n=2$ is done. check the rest
By induction

Example: $X \sim \text{Binomial}(n, p)$
$E(X) = np$.

To get this
$$E(X) = \sum_{k=0}^{n} k \, P(X=k)$$
$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Homework

Another way:

If $X \sim \text{Binomial}(n, p)$

then we can think of X as

$$X = X_1 + X_2 + X_3 + \dots + X_n$$

where X_i's are independent and identiacally distributed (i.i.d.)

Bernoulli distribution.

~~$E(X_i) =$~~

X_i	Prob.
1	p
0	$1-p$

$$E(X_i) = 1 * p + 0 * (1-p) = p$$ for all i

Therefore, using the last theorem

$$E(X) = E(X_1) + E(X_2) + \dots + E(X_n)$$

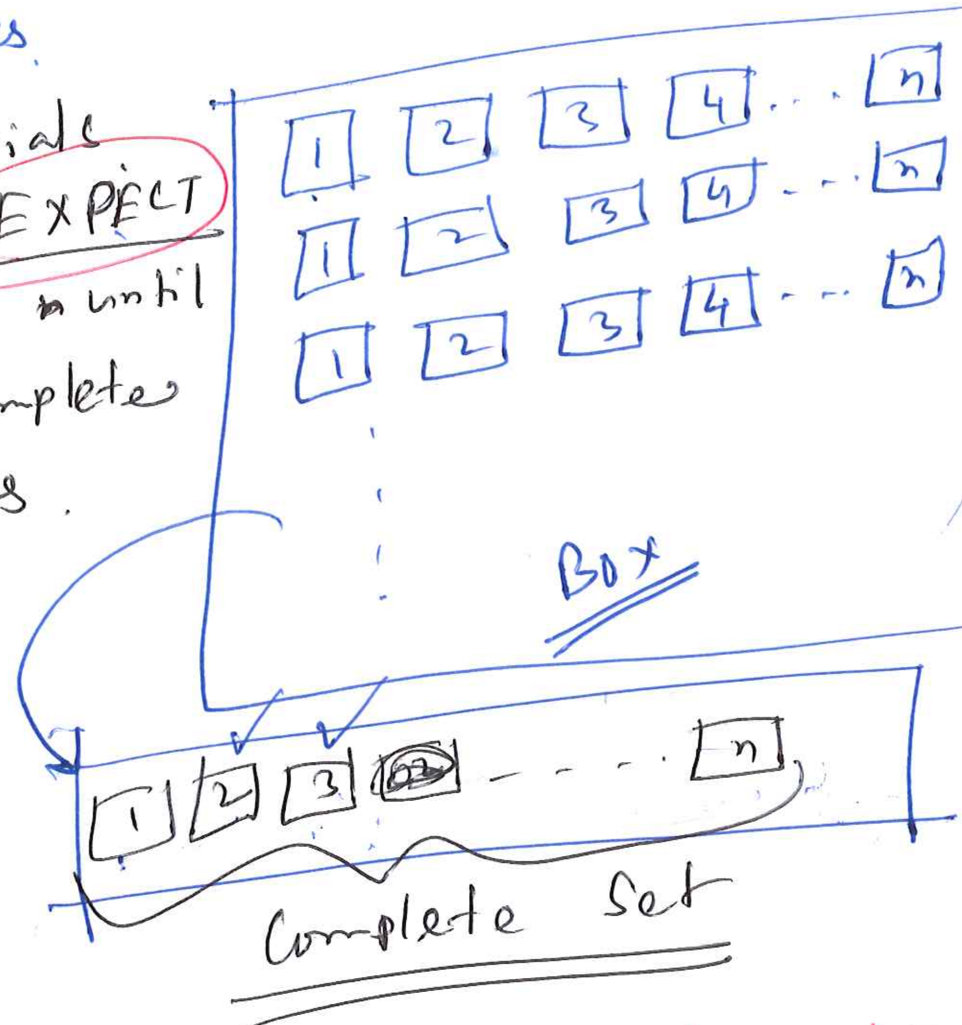
$$= \underbrace{p + p + \dots + p}_{n\text{-times}}$$

$$= \boxed{np}$$

Example: (Coupon Collection)

- Suppose that you collect coupons.
- There are n-distinct types of coupons.
- On each trial you are equally likely to get a coupon of any of the types.

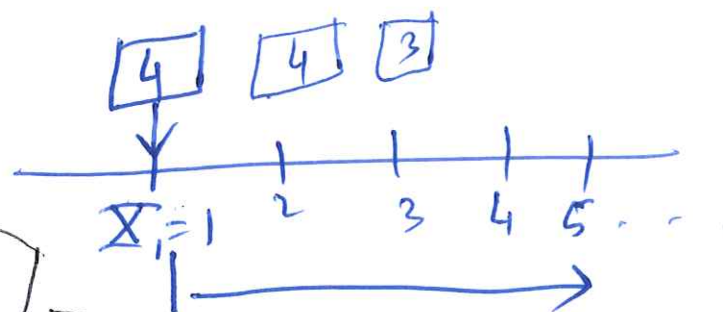
How many trials would you EXPECT to go through until you had a complete set of coupons.



Solution: X_1 : # of of trials up to and including the trial on which the first coupon is collected $X_1 = 1$

X_2: # of trials from that point (picking up the first coupon) up to and including the trial on which the next coupon DIFFERENT FROM the first is obtained.

Similarly X_3 etc.



Time taken to get the different coupon.

Then the total # of trials X

$$= X_1 + X_2 + X_3 + \cdots + X_n$$

So, $E(X) = E(X_1) + E(X_2) + E(X_3) + \cdots + E(X_n)$

$X_1 = 1$

$E(X_1) = 1$

~~$E(X_2)$?~~

~~Supp~~ Consider X_r

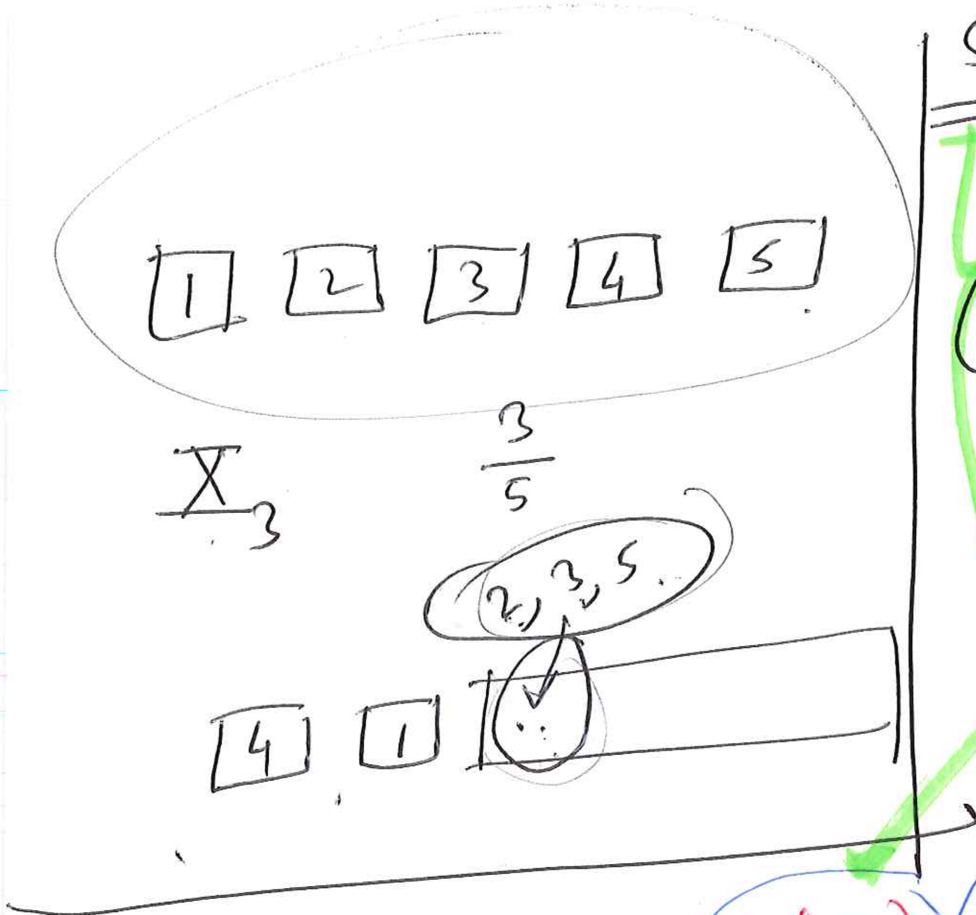
~ Geometric distribution

probability of success

$(r-1)$ coupons are picked up

$n-(r-1) = n-r+1$ coupons left

Probability of success $= \dfrac{n-r+1}{n}$



So,

$$\boxed{E(X_r) = \frac{n}{n-r+1}}$$

(By the property of geometric distribution)

Hence $E(X) = E(X_1) + E(X_2) + \dots + E(X_n)$

$$= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \dots + \frac{n}{1}$$

$$= \boxed{n \sum_{r=1}^{n} \frac{1}{r}}$$

Variance and Standard Deviation

Definition: If X is a random variable with expectation $E(X)$, then the variance of X is

$$Var(X) = E\left[\left(X - E(X)\right)^2\right]$$

Suppose X is continuous with pdf $f(x)$.

Think $Y = \left(X - E(X)\right)^2 = g(X)$

$$E(Y) = E\left(X - E(X)\right)^2$$

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

$$= \int_{-\infty}^{\infty} \left(x - E(X)\right)^2 f(x)\, dx$$

(a number) $= \mu$

def

So, $$\boxed{Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx}$$

Theorem: Suppose $Y = a + b\,X$

Then $Var(Y) = b^2\, Var(X)$.

Proof:

$$Var(Y) \overset{def}{=} E\left(\left(Y - E(Y)\right)^2\right)$$

$$= E\left(\left(\cancel{a} + b\,X - \left(\cancel{a} + b\,E(X)\right)\right)^2\right)$$

$$= E\left(b^2\left(X - E(X)\right)^2\right)$$

$$= b^2\, \boxed{E\left(X - E(X)\right)^2}$$

$$= b^2\, Var(X)$$

//

Bernoulli distribution:

X	Prob.
✓ 1	p
✓ 0	$(1-p)$

$$E(X) = 1 * p + 0\,(1-p) = p.$$

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) = E\left(\left(X - p\right)^2\right)$$

$$= (1-p)^2 \cdot P(X=1) + (0-p)^2\, P(X=0)$$

$$= (1-p)^2 \cdot p + (-p)^2 \cdot (1-p)$$

$$= (1+p^2-2p)\, p + p^2(1-p)$$

$$= p + \cancel{p^3} - 2p^2 + p^2 - \cancel{p^3}$$

$$= p - p^2 = \boxed{p(1-p)}.$$

Variance for Bernoulli distribution

Normal distribution

$X \sim N(\mu, \sigma^2)$

$E(X) = \mu$ (done last time)

$$\text{Var}(X) = E\left(\left(X - E(X)\right)^2\right)$$

$$= E\left[(X-\mu)^2\right]$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\, dx$$

$$= \int_{-\infty}^{\infty} \sigma^2 u^2 \cdot \frac{1}{\cancel{\sigma}\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \cancel{\sigma}\, du$$

Sub. $u = \frac{x-\mu}{\sigma}$

$du = \frac{dx}{\sigma}$

@ $x = -\infty$, $u = -\infty$

$x = +\infty$, $u = +\infty$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 \cdot e^{-\frac{1}{2}u^2} du$$

even

$$= \frac{\sigma^2}{\sqrt{2\pi}} \cdot 2 \int_{0}^{\infty} u^2 \cdot e^{-\frac{1}{2}u^2} du$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} (2z) \cdot e^{-z} \cdot \frac{dz}{\sqrt{2z}}$$

Sub $z = \frac{u^2}{2}$

$dz = u\, du$

$du = \frac{dz}{\sqrt{2z}}$

$$= \frac{2\sigma^2}{\cancel{\sqrt{2\pi}}} \cancel{\sqrt{2}} \int_{0}^{\infty} z^{1/2} e^{-z} dz$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} z^{\frac{3}{2}-1} e^{-z} dz$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right)$$

$$= \frac{\cancel{2}\sigma^2}{\cancel{\sqrt{\pi}}} \cdot \frac{1}{\cancel{2}}\cancel{\sqrt{\pi}}$$

$$= \boxed{\sigma^2}$$

Variance of $N(\mu, \sigma^2)$