November 8, 2016

x_1 and x_2 are random variables

$T = x_1 + x_2$

$E[T] = \sum_{t \in Supp[T]} t\,p(t)$ ← Not a good strategy.

$$> T = g(x_1, x_2) = g(\overset{\text{vector}}{\vec{x}})$$

$$E[g(x_1, x_2)] = \sum_{\vec{x} \in Supp[\vec{X}]} g(\vec{x})\,P(\vec{x})$$

$$= \sum_{\substack{(x_1, x_2) \\ \in Supp[X_1] \times Supp[X_2]}} g(x_1, x_2)\,\overbrace{P(x_1, x_2)}^{\text{Joint Mass Function (JMF)}}$$

$$= \sum_{x_1 \in Supp[X_1]} \sum_{x_2 \in Supp[X_2]} g(x_1, x_2)\,P(x_1, x_2)$$

$$E[T] = E[x_1 + x_2] = \sum_{x_1 \in Supp[X_1]} \sum_{x_2} (x_1 + x_2)\,P(x_1, x_2)$$

$$= \sum_{x_1} \sum_{x_2} \overset{\text{constant}}{x_1}\,P(x_1, x_2) + \sum_{x_2} \sum_{x_1} \overset{\text{constant}}{x_2}\,P(x_1, x_2)$$

$$= \sum_{x_1} x_1 \sum_{x_2} P(x_1, x_2) + \sum_{x_2} x_2 \sum_{x_1} P(x_1, x_2)$$

Now we're stuck... Soften the assumption.

Say x_1, x_2 are independent random variables.

$\Rightarrow P(x_1, x_2) = P(x_1) \cdot P(x_2)$

Replace JMFs with $P(x_1) \cdot P(x_2)$

$$E[T] = \sum_{x_1} x_1 \sum_{x_2} \underset{\text{constant}}{P(x_1)} \cdot P(x_2) + \sum_{x_2} x_2 \sum_{x_1} P(x_1) \cdot \underset{\text{constant}}{P(x_2)}$$

$$= \underbrace{\sum_{x_1} x_1 P(x_1)}_{=E[X_1]} \underbrace{\sum_{x_2} P(x_2)}_{=1} + \underbrace{\sum_{x_2} x_2 P(x_2)}_{=E[X_2]} \underbrace{\sum_{x_1} P(x_1)}_{=1}$$

For two random variables that are independent, the expectation of the sum of x_1, x_2 is the sum of the expectation.

> If x_1, x_2 are <u>not</u> independent random variables, we need to figure out $\sum_{x_2} P(x_1, x_2)$ and $\sum_{x_1} P(x_1, x_2)$

Suppose that the

$$\text{Supp}[X_1] = \{1, 7, 19\} \qquad \text{Supp}[X_2] = \{5, 23, 88\}$$

we use a grid to show the probability of joint events happening @ same time

JMFs $P(x_1, x_2)$

		x_1			
		1	7	19	$P(x_2)$
x_2	5	$\frac{1}{15}$	$\frac{1}{3}$	$\frac{2}{15}$	$\frac{16}{30}$
	23	$\frac{1}{30}$	$\frac{1}{10}$	$\frac{1}{30}$	$\frac{5}{30}$
	88	$\frac{1}{30}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{9}{30}$
	$P(x_1)$	$\frac{4}{30}$	$\frac{19}{30}$	$\frac{7}{30}$	1

Marginal PMFs, $P(x_2)$

Rule: $\sum_{x_1} \sum_{x_2} P(x_1, x_2) = 1$

↑

Generalization from the rule we have for r.v.: $\sum_{x \in \text{Supp}[X]} p(x) = 1$

Marginal PMFs, $P(x_1)$

$$P(X_1 = 1) = P(X_1 = 1, X_2 = 5) + P(X_1 = 1, X_2 = 23) + P(X_1 = 1, X_2 = 88)$$

Remember: Law of Total Probability.

$$P(x_1) = \sum_{x_2} P(x_1, x_2)$$

But where did the x_2 go? It was margined out. x_2 is a "marginal error". Similar to finding $g(x) = \int f(x,y)\,dy$. Where did the y go? It got "integrated out".

> Back to finding $E[x_1 + x_2]$ ...

$$E[T] = \underbrace{\sum_{x_1} x_1 P(x_1)}_{= E[x_1]} + \underbrace{\sum_{x_2} x_2 P(x_2)}_{= E[x_2]}$$

$$E[T] = E[X_1 + X_2] = E[X_1] + E[X_2]!$$

> Are x_1 and x_2 independent?

$$P(1, 88) \overset{?}{\neq} P(1)\,P(88)$$

$$\frac{1}{30} \neq \left(\frac{4}{30} \cdot \frac{9}{30} = \frac{36}{300}\right)$$

No, x_1 and x_2 are <u>dependent</u>.

- General Rule:

  $X_1, X_2, \ldots, X_n$ r.v's

  $\Rightarrow E[T] = E[X_1, X_2, \ldots, X_n]$

  $= E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = E[X_1] + \ldots + E[X_n]$

  or $= \mu_1 + \ldots + \mu_n$

- Let's say $X_1, \ldots, X_n$ are identically distributed (but not necessarily independent). What is the $E[T]$?

  $E[T] = \sum_{i=1}^{n} E[X_i] = E[X_1 + \ldots + X_n] = n\mu$

  (because if they are identically distributed, each μ is the same, which means their PMFs are the same, which means that their centerpoints are the same.)

- $X \sim \text{Binom}(n, p)$

  $X = X_1 + \ldots + X_n$

  $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bern}(p)$

  $E[X] = np$ (We proved this by using Binom def. We can also use the $E[X]$ for Bern. There are n r.v.'s in a Bern iid, and one has an $E[X] = p$. Since Binom is a sum of all the n Bern r.v.'s, the Binom $E[X]$ is $n \cdot p$.)

- $X \sim \text{NegBin}(r, p)$

  $E[X] = \sum_{x=r}^{\infty} x \binom{x-1}{r-1} (1-p)^{x-r} p^r$

  Recall that $X_1, X_2, \ldots, X_r \overset{iid}{\sim} \text{Geom}(p)$.

  $X = X_1 + X_2 + \ldots + X_r \sim \text{NegBin}(r, p)$.

  $$E[T] = \underset{\uparrow}{r\mu} = r \cdot \underset{\uparrow}{\frac{1}{p}} = \frac{r}{p}$$

  b/c it is an iid, it is identically distributed.

  b/c the E[X] for binom $= \frac{1}{p}$.

> $X \sim \text{Hyper}(n, K, N)$

$$E[X] = \sum_{x \in \text{Supp}[X]} x \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

Is there a way to simplify this when there are so many base cases? How can we do this generally?

$X \sim \text{Hyper}(1, K, N) = \text{Bern}\left(\frac{K}{N}\right)$

$\overset{\downarrow}{X} = X_1 + X_2 + \dots + X_n$

s.t. $X_1, X_2, \dots, X_n$ are identically distributed Bern$\left(\frac{K}{N}\right)$, but <u>NOT</u> independent. (B/c say you have N coins, K=1 with spots. If you don't look at the coins all at once, but one by one, they all have a probability of $\frac{K}{N}$ of having a spot, <u>but once you get a coin with a spot, the rest</u> of the coins have a probability of 0 of having a spot = Dependent)

$$E[X] = \underset{\substack{\uparrow \\ \text{b/c they} \\ \text{are identically} \\ \text{distributed}}}{n\mu} = n \underset{\substack{\uparrow \\ \text{b/c } E[X] \text{ of} \\ \text{Bern is } p = \frac{K}{N}.}}{\frac{K}{N}}$$

> Review of Variance

$$\begin{aligned}
\text{Var}[X] &:= E[(X-\mu)^2] \\
&= E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] + E[-2\mu X] + E[\mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2
\end{aligned}$$

$\text{Var}[X] = E[X^2] - \mu^2$

$\sigma^2 = E[X^2] - \mu^2$

$E[X^2] = \sigma^2 + \mu^2$

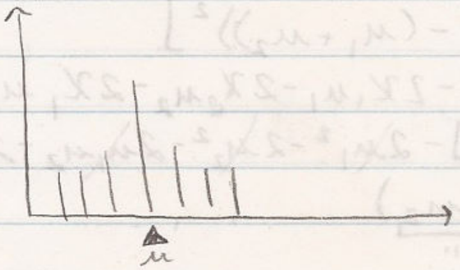
Not on exam

$E[X^2]$ - "second moment" $E[X^k]$ - "k^{th} moment"

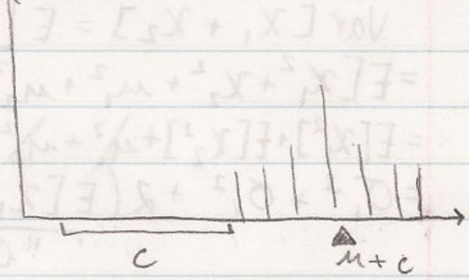
$E[(X-\mu)]$ - "first centered moment" $E[|X-\mu|^k]$ - "skewness"

$E\left[\frac{(X-\mu)^2}{\sigma^2}\right]$ - "third centered moment" $E\left[\frac{|X-\mu|^4}{\sigma^4}\right]$ - "Kurtosis"

Var[X]



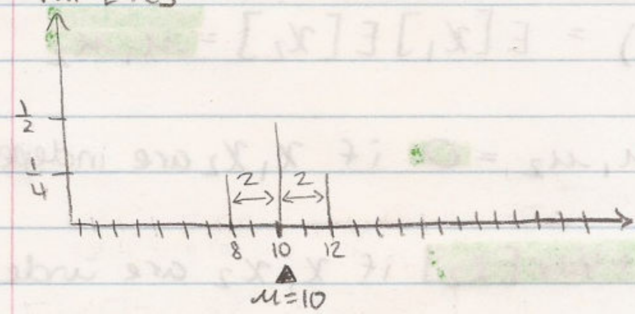
Var[X+c] s.t. $c \in \mathbb{R}$



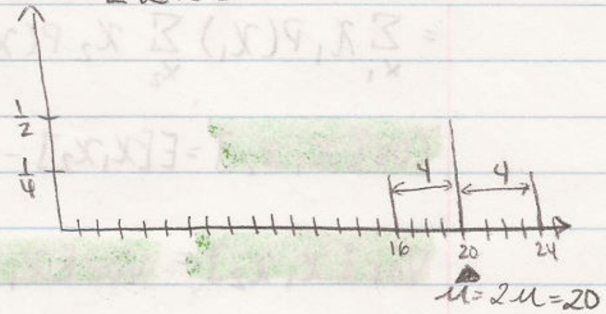
- Distances from the balance point do not change.

$$Var[X+c] = E[((X+c)-(\mu+c))^2]$$
$$= E[(X-\mu)^2]$$
$$Var[X+c] = Var[X].$$

Var[X]



Var [2X]



$$Var[X] = 2^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{2} = \frac{1}{4}$$

$$Var[2X] = 4^2 \cdot \frac{1}{2} + 4^2 \cdot \frac{1}{2} = 16$$
$$= 4 \cdot Var[X]$$
$$= 2^2\, Var[X]$$

$$Var[aX] = E[(aX - a\mu)^2] = E[(a(X-\mu))^2] - E[a^2(X-\mu)^2]$$
$$= a^2 E[(X-\mu)^2] = a^2\, Var[X] \text{ or } a^2 \sigma$$

$$SE[aX] = \sqrt{Var[aX]} = \sqrt{a^2\sigma^2} = |a|\sigma$$
$$Var[aX+c] = a^2\sigma^2$$
$$SE[aX+c] = \sqrt{Var[aX+c]} = \sqrt{Var[aX]} = |a|\sigma$$

x_1, x_2 are r.v.'s

$$Var[x_1 + x_2] = E[((x_1+x_2) - (\mu_1+\mu_2))^2]$$

$$= E[x_1^2 + x_2^2 + \mu_1^2 + \mu_2^2 + 2x_1x_2 - 2x_1\mu_1 - 2x_2\mu_2 - 2x_1\mu_2 - 2x_2\mu_1 + 2\mu_1\mu_2$$

$$= E[x_1^2] + E[x_2^2] + \cancel{\mu_1^2} + \cancel{\mu_2^2} + 2E[x_1x_2] - \cancel{2}\mu_1^2 - \cancel{2}\mu_2^2 - \cancel{2\mu_1\mu_2} - 2\mu_2\mu_1 + \cancel{2\mu_1\mu_2}$$

$$= \sigma_1^2 + \sigma_2^2 + 2(\underbrace{E[x_1x_2] - \mu_1\mu_2}_{\text{"Covariance"}})$$

$$Cov[x_1, x_2] := E[x_1x_2 - \mu_1\mu_2]$$

If x_1, x_2 are independent but not necessarily identically distributed, what is $E[x_1x_2]$?

$$E[x_1x_2] = \sum_{x_1}\sum_{x_2} x_1x_2 P(x_1, x_2) = \sum_{x_1}\sum_{x_2} x_1, x_2 P(x_1) P(x_2)$$

$$= \sum_{x_1} x_1 P(x_1) \sum_{x_2} x_2 P(x_2) = E[x_1] E[x_2] = \mu_1\mu_2$$

$Cov[x_1, x_2] = E[x_1x_2] - \mu_1\mu_2 = 0$ if $x_1 x_2$ are independent.

$Var[x_1 x_2] = Var[x_1] + Var[x_2]$ if x_1, x_2 are independent.

General Rule (if $x_1, \ldots, x_n$ are independent)

$$Var[T] = Var[x_1 + \ldots + x_n] = Var[x_1] + \ldots + Var[x_n]$$

$$= \sum_{i=1}^{n} Var[x_i] = \sum_{i=1}^{n} \sigma_i^2$$

If $x_1, x_2, \ldots, x_n \overset{iid}{\sim}$

$$Var[T] = n\sigma^2$$

$$Var[\bar{x}] = Var\left[\frac{T}{n}\right] = \frac{1}{n^2} Var[T] = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

$$SE[\bar{x}] = \frac{\sigma}{\sqrt{n}}$$