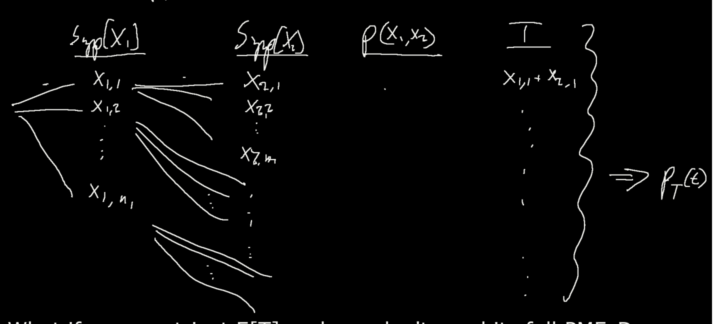
$T = X_1 + X_2 \sim p_T(t)$



What if you want just E[T] and you don't need its full PMF. Do you need its PMF? We know E[T] = $\sum_{t \in Supp[T]}$ t p(t) using the definition formula.

It turns our there's an easier way to calculate this expectation without requiring the entire PMF.

$T = X_1 + X_2 = g(X_1, X_2)$

$$E[T] = \underbrace{E[g(X_1, X_2)] = \sum_{x_1 \in Supp[X_1]} \sum_{x_2 \in Supp[X_2]} g(x_1, x_2)\, p(x_1, x_2)}_{\text{general formula}}$$

$\underbrace{JMF}$ joint mass function

$\sum 7 c_i = 7 \sum c_i$

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

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

If $X_1, X_2 \overset{ind}{\sim} \Rightarrow p(x_1, x_2) = p(x_1)\, p(x_2)$

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

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

If they're not independent we need to simplify $\sum_{x_1}$ p(x_1, x_2) and $\sum_{x_2}$ p(x_1, x_2). Let's talk more about JMFs

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

$p(x_1, x_2)$

X_2 \ X_1	1	7	19	
5	$\frac{1}{15}$	$\frac{1}{3}$	$\frac{2}{15}$	$\frac{16}{30} = p(X_2 = 5)$
23	$\frac{1}{30}$	$\frac{1}{10}$	$\frac{1}{30}$	$\frac{5}{30} = p(X_2 = 23)$
88	$\frac{1}{30}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{9}{30} = p(X_2 = 88)$
	$\frac{4}{30}$	$\frac{19}{30}$	$\frac{7}{30}$	1

$\underbrace{\quad}_{p(x_1)}$; $p(x_2)$

marginal probability

marginal PMF, derived from the JMF by "margin-ing out X_1"

$$\Rightarrow \sum_{x_1} p(x_1, x_2) = p(x_2), \quad \sum_{x_2} p(x_1, x_2) = p(x_1)$$

$\frac{p(X_1 = 1, X_2 = 5)}{p(X_2 = 5)}$

$X_1, X_2 \overset{ind}{\sim} ?$ $\quad p(X_1 = 1) \overset{?}{=} p(X_1 = 1 \mid X_2 = 5)$

$\frac{2}{15} = \frac{4}{30} \overset{?}{=} \frac{\frac{1}{15}}{\frac{16}{30}} = \frac{2}{16}$ NO!

X_1, X_2 are dependent

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

$$= \sum_{x_1} x_1 p(x_1) + \sum_{x_2} x_2 p(x_2) = \boxed{E[X_1] + E[X_2] = E[T]}$$

ALWAYS!!!!

For any discrete rv's $X_1, X_2, \ldots, X_n$, $T = X_1 + \ldots + X_n$, $\bar{X}_n = \frac{T}{n}$

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

If the rv's are identically distributed

$$E[\bar{X}] = E\left[\frac{T}{n}\right] = \frac{1}{n} E[T] = \frac{n\mu}{n} = \mu$$

X ~ Hyper(n, K, N) := $\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$, $E[X] = \sum_{x \in Supp[X]} x \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$

this is really hard to do!

To solve for the expectation easily, we conceptualize X as a sum of random variables $B_1, B_2, \ldots, B_n$. Each B_i is zero-one and represents if you got a special ball on the ith draw.

$$E[X] = \sum_{i=1}^{n} E[B_i] = \boxed{n \frac{K}{N}}$$

What is the distribution of the B's? They must be Bernoulli:

$B_1 \sim Bern\left(\frac{K}{N}\right)$, $E[B_1] = \frac{K}{N}$

$B_2 \sim Bern\left(\frac{K}{N}\right)$

$\vdots$

$B_n \sim Bern\left(\frac{K}{N}\right)$

These B's are identically distributed.

Are they independent? NO

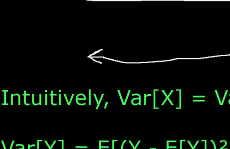
Let's return to variance. Let's derive a useful formula:

Var[X] := E[(X - μ)²] = E[X² - 2Xμ + μ²] = E[X²] + E[-2Xμ] + E[μ²]
= E[X²] -2μE[X] + μ² = E[X²] -2μ² + μ² = E[X²] - μ²

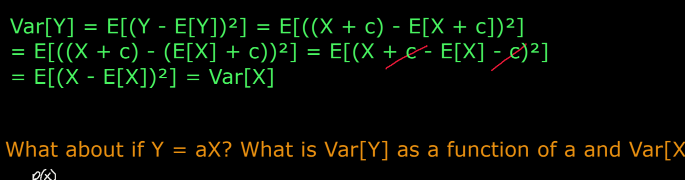
⇒ Var[X] = E[X²] - μ²
⇒ E[X²] = Var[X] + μ² = σ² + μ²

Note: E[X] is called the "first moment",
E[X²] is called the "second moment",
E[X³] is called the "third moment",
...
E[X^k] is called the "kth moment"

Recall Y = aX + c where a, c ∈ ℝ constants. We derived that E[Y] = aE[X] + c. What about Var[Y]?



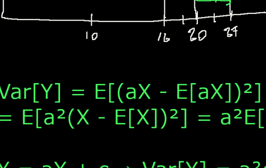
$Y = X + c, \quad c > 0$



Intuitively, Var[X] = Var[X + c]. Let's prove it for Y = X + c.

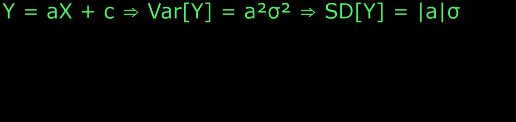
Var[Y] = E[(Y - E[Y])²] = E[((X + c) - E[X + c])²]
= E[((X + c) - (E[X] + c))²] = E[(X + ~~c~~ - E[X] - ~~c~~)²]
= E[(X - E[X])²] = Var[X]

What about if Y = aX? What is Var[Y] as a function of a and Var[X]?



$Y = 2X$

expansion



Var[Y] = E[(aX - E[aX])²] = E[(aX - aE[X])²] = E[(a(X - E[X]))²]]
= E[a²(X - E[X])²] = a²E[(X - E[X])²] = a²Var[X] = a²σ²

Y = aX + c ⇒ Var[Y] = a²σ² ⇒ SD[Y] = |a|σ