$T = X_1 + X_2 \Rightarrow E[T] = \mu_1 + \mu_2$ always $Var[T]$?

$$Var[X_1+X_2] = E\left[\left((X_1+X_2) - E[X_1+X_2]\right)^2\right] = E\left[(X_1+X_2-\mu_1-\mu_2)^2\right]$$

$$= E\left[X_1^2 + X_2^2 + \mu_1^2 + \mu_2^2 + 2X_1X_2 - 2X_1\mu_1 - 2X_1\mu_2 - 2X_2\mu_1 - 2X_2\mu_2 + 2\mu_1\mu_2\right]$$

$$= E[X_1^2] + E[X_2^2] + \cancel{\mu_1^2} + \cancel{\mu_2^2} + 2E[X_1X_2] - 2\mu_1 \underbrace{E[X_1]}_{\mu_1} - 2\mu_2 \underbrace{E[X_1]}_{\mu_1} - 2\mu_1 \underbrace{E[X_2]}_{\mu_2} - 2\mu_2 \underbrace{E[X_2]}_{\mu_2}$$

$-2\mu_1^2 \quad -2\mu_2\mu_1 \quad \cancel{-2\mu_1\mu_2} \quad -2\mu_2^2 \quad \cancel{+2\mu_1\mu_2}$

$$= E[X_1^2] + E[X_2^2] + 2E[X_1X_2] - \mu_1^2 - \mu_2^2 - 2\mu_1\mu_2$$

$E[X^2] = \sigma^2 + \mu^2$ (last class)

$$= \sigma_1^2 + \cancel{\mu_1^2} + \sigma_2^2 + \cancel{\mu_2^2} + 2E[X_1X_2] - \cancel{\mu_1^2} - \cancel{\mu_2^2} - 2\mu_1\mu_2$$

$$= \sigma_1^2 + \sigma_2^2 + 2\left(E[X_1X_2] - \mu_1\mu_2\right)$$

Define the covariance of two rv's:

$$Cov[X_1, X_2] := E[X_1X_2] - \mu_1\mu_2$$

$$= \sigma_1^2 + \sigma_2^2 + 2\,Cov[X_1, X_2]$$

$$= \sigma_1^2 + \sigma_2^2 \Rightarrow Var[X_1+X_2] = Var[X_1] + Var[X_2] \text{ if } X_1, X_2 \overset{ind}{\sim}$$

$X_1, X_2 \overset{ind}{\sim}$

$$E[X_1X_2] = \sum_{x_1 \in Supp[X_1]} \sum_{x_2 \in Supp[X_2]} x_1 x_2\, p_{X_1,X_2}(x_1, x_2)$$

this cannot be simplified

$X_1, X_2 \overset{ind}{\sim}$

$$= \sum_{x_1} \sum_{x_2} x_1 x_2\, p_{X_1}(x_1)\, p_{X_2}(x_2) = \sum_{x_1} x_1 p_{X_1}(x_1) \sum_{x_2} x_2 p_{X_2}(x_2) = E[X_1]E[X_2] = \mu_1\mu_2$$

$$Cov[X_1, X_2] = E[X_1X_2] - \mu_1\mu_2 = \mu_1\mu_2 - \mu_1\mu_2 = 0$$

Covariance is zero if the two rv's are independent

If $X_1, \ldots, X_n \overset{ind}{\sim}$, $T := \sum_{i=1}^n X_i$, $\bar{X} := \frac{T}{n}$

$$Var[T] = \sum_{i=1}^n \sigma_i^2 = n\sigma^2 \quad \text{if } X_1, \ldots, X_n \overset{iid}{\sim}$$

$$Var[\bar{X}] = Var\left[\frac{T}{n}\right] = \frac{1}{n^2} Var[T] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0$$

$T \sim Binom(n, p)$

$$Var[T] = E[(T-\mu)^2] = E[(T-np)^2] = \sum_{t=0}^n (t-np)^2 \binom{n}{t} p^t (1-p)^{n-t} \quad \text{HARD!!}$$

Recall $X_1, \ldots, X_n \overset{iid}{\sim} Bern(p) \Rightarrow T = X_1 + \ldots + X_n \sim Bin(n, p)$

$$Var[T] = Var[X_1 + \ldots + X_n] = n\sigma^2 = \boxed{np(1-p)}$$

$Supp[Y] = \mathbb{N}$

$$Y \sim Geom(p) := (1-p)^{y-1}p, \quad E[Y] = \frac{1}{p}, \quad Var[Y] = E[Y^2] - \mu^2 = E[Y^2] - \frac{1}{p^2}$$

let $z = y-1 \Rightarrow y = z+1$

$$E[Y^2] = \sum_{y=1}^{\infty} y^2 (1-p)^{y-1} p = p \sum_{z=0}^{\infty} (z+1)^2 (1-p)^z$$

$$= p \sum_{z=0}^{\infty} (z^2+2z+1)(1-p)^z = (1-p)\underbrace{\sum_{z=1}^{\infty} z^2 (1-p)^{z-1} p}_{E[Y^2]} + 2(1-p)\underbrace{\sum_{z=1}^{\infty} z(1-p)^{z-1} p}_{E[Y] = \frac{1}{p}} + \cancel{p}\underbrace{\sum_{z=0}^{\infty} (1-p)^z}_{\frac{1}{1-(1-p)} = \frac{1}{\cancel{p}}}$$

Geometric Series

$$\Rightarrow E[Y^2] = (1-p)E[Y^2] + \frac{2(1-p)}{p} + 1$$

$$\Rightarrow E[Y^2] - (1-p)E[Y^2] = \frac{2-2p}{p} + \frac{p}{p} = \frac{2-p}{p}$$

$$\Rightarrow E[Y^2]\underbrace{(1-(1-p))}_{p} = \frac{2-p}{p} \Rightarrow E[Y^2] = \frac{2-p}{p^2}$$

$$Var[Y] = E[Y^2] - \frac{1}{p^2} = \frac{2-p}{p^2} - \frac{1}{p^2} = \boxed{\frac{1-p}{p^2}}$$

Here is a really interesting fact about the geometric rv

$X_1, X_2, \ldots \overset{iid}{\sim} Bern(p)$, $X \sim Geom(p)$ when did this sequence stop / get a first 1?

0000000000 _ _ _ _ _ _ _ _

$$P(X=17 \mid X>10) = \frac{P(X=17 \text{ and } X>10)}{P(X>10)}$$

def of cond. prob.

$$= \frac{P(X=17)}{P(X>10)} = \frac{p(17)}{1-F(10)} =$$

$$= \frac{(1-p)^{17-1}p}{(1-p)^{10}} = (1-p)^{7-1}p = p(7)$$

$a, b > 0$

$$P(X = a+b \mid X > a) = \frac{P(X=a+b \text{ and } X>a)}{P(X>a)} = \frac{p(a+b)}{1-F(a)}$$

$$= \frac{(1-p)^{a+b-1}p}{(1-p)^a} = \frac{\cancel{(1-p)^a}(1-p)^{b-1}p}{\cancel{(1-p)^a}}$$

$$= (1-p)^{b-1}p = P(X=b)$$

This fact is called the "memorylessness" property of the geometric rv. See the movie Memento!

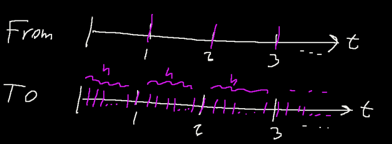
Midterm II ↑

Final exam ↓

$$T \sim Geom(p) := \overbrace{(1-p)^{t-1}p}^{p(t)}, \quad F(t) = 1-(1-p)^t$$

Consider the unit here is seconds. So if t = 7, you got the first "success" on the 7th second and 7th experiment.

Imagine within each second there are n Bernoulli experiments occuring:



If we want a new rv whose unit is time (not # of experiments) we can write its PMF and CDF as:

$$T_n \sim Geom(p) \overset{new}{:=} (1-p)^{nt-1}p = p(t), \quad F(t) = 1-(1-p)^{nt}$$

#of experiments done by time t

If n → ∞, then T → Deg(0) because I'm running bajillions of experiments in every tiny time period.

How do we make this scenario interesting? Let p → 0 as n → ∞ and we force λ := np ⇒ p = λ / n

$$T_n \sim Geom(p) := \left(1-\frac{\lambda}{n}\right)^{nt-1}\frac{\lambda}{n} = p(t), \quad F(t) = 1-\left(1-\frac{\lambda}{n}\right)^{nt}$$

$$\lim_{n\to\infty} p(t) = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{nt-1}\frac{\lambda}{n} = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{nt-1} \underbrace{\lim_{n\to\infty}\frac{\lambda}{n}}_{0} = 0 \quad \text{No PMF!}$$

$\lim f(x)g(x) = \lim f(x) \lim g(x)$

The probability that this new rv is t is always 0 no matter what t.

$\lim c + f(x) = c + \lim f(x)$ $\lim f(x)^a = (\lim f(x))^a$

$$\lim_{n\to\infty} F(t) = \lim_{n\to\infty} 1 - \left(1-\frac{\lambda}{n}\right)^{nt} = 1 - \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{nt} = 1 - \left(\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\right)^t$$

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = e^a$$

$$= 1 - e^{-\lambda t}$$

$$F(t) := P(T \le t)$$

Is this a valid CDF? Let's verify the three properties of the CDF:

(1) $\lim_{t\to-\infty}$ F(t) = 0 $= P(T=0) = 0$ ✓

$\lambda = np > 0$
$n > 0,\ p \in (0, 1)$

(2) $\lim_{t\to\infty}$ F(t) = 1 $\lim_{t\to\infty} 1 - e^{-\lambda t} = 1 - \lim_{t\to\infty}\frac{1}{e^{\lambda t}} = 1 - 0 = 1$ ✓

(3) F(t) is monotonically increasing

$$\frac{d}{dt}[F(t)] = \frac{d}{dt}\left[1-e^{-\lambda t}\right] = --\lambda e^{-\lambda t} = \lambda e^{-\lambda t} > 0 \checkmark$$

Thus... F(t) is a valid CDF. So we have a rv! It just doesn't have a PMF! This is our first "continous rv". Continuous rv's do not have PMF's. They have CDFs and they have PDFs (probability density distributions). Why are they called continuous? Let T denote the rv we found as n → ∞.

$$Supp[T] = (0, \infty), \quad |Supp[T]| = |\mathbb{R}|$$

the "continuum" or "uncountable infinity"

This is no longer a "discrete rv" but a "continuous rv".

$p(3) = P(T=3) = 0$ this is zero for every t!!!

$$P(T \in [2.995, 3.005])$$

$$= P(T \le 3.005) - P(T \le 2.995)$$

$$= F(3.005) - F(2.995) = \left(1-e^{-\lambda 3.005}\right) - \left(1-e^{-\lambda 2.995}\right) \ne 0$$

$a < b$

⇒ For a continuous rv T, P(T ∈ [a, b]) = F(b) - F(a) $= \int_a^b f(t)\,dt$

$$f(t) := \frac{d}{dt}\left[F(t)\right]$$

By the Fundamental Thm of Calculus

This is the definition of the "probability density function" (PDF). Continuous rv's have a CDF and PDF but not a PMF!