Suppose that

$$X_1, X_2, \ldots, X_n \stackrel{iid}{\sim} \text{Bernoulli}(p)$$

Question:

What is the distribution of $Y = \sum_{i=1}^{n} X_i$?

- Each X_i takes on the value 1("success") with probability p and 0 ("failure") with probability 1-p.
- Summing them up will give the total number of 1's which is the total number of "S"s in n trials.

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$$

$$\Downarrow$$

$$Y = \sum_{i=1}^{n} X_i \sim \text{bin}(n, p)$$

Not all random variables are so easily interpreted...

Moment Generating Functions:

Let X be a random variable.

It's **moment generating function** (mgf) is denoted and defined as

$$M_X(t) = E[e^{tX}]$$

It can be used to produce "moments" for X: $E[X], E[X^2], E[X^3], \ldots$

Moment generating functions also uniquely identify distributions.

$$\boxed{M_X(t)} = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \boxed{f_X(x)} dx$$

Example: $X \sim \text{Bernoulli}(p)$

$$M_X(t) = E[e^{tX}] = \sum_x e^{tx} f_X(x)$$

$$= \sum_x e^{tx} P(X = x)$$

$$\begin{aligned} M_X(t) &= \sum_x e^{tx} P(X = x) \\ &= e^{t\cdot 0} P(X = 0) + e^{t\cdot 1} P(X = 1) \\ &= 1 \cdot (1 - p) + e^t \cdot p \\ &= 1 - p + pe^t \end{aligned}$$

Example: $X \sim \text{bin}(n, p)$

$$\begin{aligned} M_X(t) &= \sum_x e^{tx} P(X = x) \\ &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \end{aligned}$$

Example: $X \sim \text{bin}(n, p)$

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}$$

Binomial Theorem:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

$$\Rightarrow \quad M_X(t) = (pe^t + 1 - p)^n$$

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from some distribution with mgf $M_X(t)$.

$$\text{Let} \quad Y = \sum_{i=1}^{n} X_i$$

The mgf for Y is

$$M_Y(t) = E[e^{tY}] = E\left[e^{t\sum_{i=1}^{n} X_i}\right]$$

$$= E\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} E[e^{tX_i}] = \left(M_{X_1}(t)\right)^n$$

all the same

by independence

$$X_1, X_2, \ldots, X_n \text{ iid}$$

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

$$\Downarrow$$

$$M_Y(t) = [M_{X_1}(t)]^n$$

Example:

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$$

Find the distribution of $Y = \sum_{i=1}^{n} X_i$.

$$M_{X_1}(t) = 1 - p + pe^t$$

$$M_Y(t) \overset{iid}{=} [M_{X_1}(t)]^n = [1 - p + pe^t]^n$$

but this is the mgf of the bin(n,p) distribution...

and this is enough to say that

$$Y \sim \text{bin}(n, p)$$

Example:

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} \exp(\text{rate} = \lambda)$$

Find the mgf for $\quad Y = \sum_{i=1}^{n} X_i$

$$M_{X_1}(t) = E[e^{tX_1}] = \int_{-\infty}^{\infty} e^{tx} f_{X_1}(x)\, dx$$

$$= \int_0^{\infty} e^{tx} \lambda e^{-\lambda x}\, dx = \int_0^{\infty} \underbrace{\lambda e^{-(\lambda - t)x}}_{\text{like an } \exp(\text{rate}=\lambda\text{-t}) \text{ pdf}}\, dx$$

$$= \frac{\lambda}{\lambda - t} \underbrace{\int_0^{\infty} (\lambda - t) e^{-(\lambda - t)x}\, dx}_{=1} = \frac{\lambda}{\lambda - t} \quad \text{for } \lambda < t$$

Example:

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \Gamma(\alpha, \beta)$$

$$M_{X_1}(t) = E[e^{tX_1}] = \int_{-\infty}^{\infty} e^{tx} f_{X_1}(x)\, dx$$

$$= \int_0^{\infty} e^{tx} \frac{1}{\Gamma(\alpha)} \beta^{\alpha} x^{\alpha-1} e^{-\beta x}\, dx$$

$$= \frac{1}{\Gamma(\alpha)} \beta^{\alpha} \int_0^{\infty} \underbrace{x^{\alpha-1} e^{-(\beta-t)x}}_{\substack{\text{like a} \\ \Gamma(\alpha, \beta - t) \\ \text{pdf}}}\, dx$$

$$= \frac{\beta^{\alpha}}{(\beta - t)^{\alpha}} \int_0^{\infty} \frac{1}{\Gamma(\alpha)} (\beta - t)^{\alpha} x^{\alpha-1} e^{-(\beta-t)x}\, dx$$

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \Gamma(\alpha, \beta)$$

$$\Rightarrow \quad M_{X_1}(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}$$

Example:

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \exp(\text{rate} = \lambda)$$

Find the distribution of $\quad Y = \sum_{i=1}^{n} X_i$.

$$M_{X_1}(t) = \frac{\lambda}{\lambda - t}$$

$$M_Y(t) \overset{\text{iid}}{=} [M_{X_1}(t)]^n = \left(\frac{\lambda}{\lambda - t}\right)^n$$

This is the mgf for the $\Gamma(n, \lambda)$ distribution.

The sum of n iid exponential rate λ random variables has a gamma distribution with parameters n and λ.

- sum of n iid Bernoulli(p) random variables is bin(n,p)

- sum of n iid exp(rate=λ) random variables is $\Gamma(\mathrm{n}, \lambda)$

- sum of m iid bin(n,p) is bin(nm,p)

- sum of n iid $\Gamma(\alpha, \beta)$ is $\Gamma(\mathrm{n}\alpha, \beta)$

- sum of n iid $N(\mu, \sigma^2)$ is $N(n\mu, n\sigma^2)$

- sum of n independent normal random variable with $X_i \sim N(\mu_i, \sigma_i^2)$ is $N(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2)$

- $X_1, X_2, \ldots, X_n$ independent with $X_i \sim N(\mu_i, \sigma_i^2)$.

$$Y = \sum_{i=1}^{n} a_i X_i \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2)$$