Recall

- Conditional variance
- Law of total variance

Then we introduced moment generating function $M_X(s)$ of a rv X

$$M_X(s) = \mathbb{E}\left(e^{sX}\right)$$

Saw a caeple of examples

i.e

Gaussian (μ, σ^2) has $M_X(s) = e^{\mu s + \frac{\sigma^2 s^2}{2}}$

Reason for its name? $\mathbb{E}(X^k) = \frac{d^k}{ds^k} M_X(s)\Big|_{s=0}$

Turns out: $M_X(s)$ determines $f_X(x)$ (or $p_X(x)$) completely*

i.e there's an inversion formula

$$M_X(s) \rightsquigarrow f_X(x) \text{ or } p_X(x)$$

we don't really use this. Insteda use a table look-up.

Example

Say X_1, X_2 are Gaussian (μ_1, σ_1^2) and (μ_2, σ_2^2) and independent.

Have already seen

$$Y = X_1 + X_2$$

is also Gaussian.

Can do this by

$$f_Y(y) = f_{X_1}(x) * f_{X_2}(x)$$

Alternative argument based on moment generating functions:

Based on the fact that when X_1, X_2 independent ; $Y = X_1 + X_2$

$$M_Y(s) = M_{X_1}(s)\, M_{X_2}(s)$$

Reason: $M_Y(s) = \mathbb{E}(s^Y) = \mathbb{E}(e^{s(X_1+X_2)}) = \mathbb{E}(e^{sX_1} e^{sX_2}) = \mathbb{E}(e^{sX_1})\,\mathbb{E}(e^{sX_2})$

Convolution in pdf $\longleftrightarrow$ Multiplication in $M_X(s)$

Here,

$$M_{X_1}(s) = e^{\mu_1 s + \frac{\sigma_1^2 s^2}{2}} \qquad M_{X_2}(s) = e^{\mu_2 s + \frac{\sigma_2^2 s}{2}}$$

$$M_Y(s) = M_{X_1}(s)\, M_{X_2}(s) = e^{(\mu_1+\mu_2)s + \frac{(\sigma_1^2+\sigma_2^2)}{2} s^2}$$

recognize as MGF of Gaussian $(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

Another context where **$M_X(s)$'s help: the sum of a random # of identically distributed (iid) rv's**

Have N, a positive integer-valued rv;

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

where

- X_k all have same pmf $p_X(x)$ OR pdf $f_X(x)$
- X_k independent of eachother and of N

First find $\mathbb{E}(Y)$ and $\mathrm{Var}(Y)$

$$\mathbb{E}(Y \mid N=n) = \sum_{k=1}^{n} \mathbb{E}(X_k) = n\mathbb{E}(X) \Rightarrow \mathbb{E}(Y \mid N) = N\mathbb{E}(X)$$

Thus by law of iterated expectations

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid N)) = \mathbb{E}(\mathbb{E}(X)N) = \mathbb{E}(X)\mathbb{E}(N)$$

As for $\mathrm{Var}(Y)$, note first that

$$\mathrm{Var}(Y \mid N=n) \overset{\text{by independence}}{=} n\,\mathrm{Var}(X)$$

Thus

$$\mathrm{Var}(Y \mid N) = N\,\mathrm{Var}(X)$$

By law of total variance

$$\mathrm{Var}(Y) = \mathbb{E}(\mathrm{Var}(Y \mid N)) + \mathrm{Var}(\mathbb{E}(Y \mid N))$$

$$= \mathrm{Var}(X)\mathbb{E}(N) + (\mathbb{E}(X))^2\,\mathrm{Var}(N)$$

Now how do we find pdf $f_Y(y)$ or pmf $p_Y(y)$

How **NOT** to do it: Say X continuous; $f_X(x)$ ← This works but not a neat way to do this

Given $N=n$, $Y = X_1 + X_2 + \cdots + X_n$ has pdf $f_X(x) * f_X(x) * \cdots * f_X(x)$

Thus

$$f_{Y|N} = f_X(x) * \cdots f_X(x) \quad n \text{ times}$$

By Law of Total Probability

$$f_Y(y) = \sum_{n=1}^{+\infty} f_{Y|N}(y|n)\, p_N(n)$$

A **cooler** way to do it: Figure out $M_Y(s)$.

$$M_Y(s) = \mathbb{E}(e^{sY}) = \mathbb{E}\left(\mathbb{E}(e^{sY} | N)\right)$$

$$\mathbb{E}(e^{sY} | N=n) = \mathbb{E}(e^{s(X_1 + \cdots + X_n)} | N=n) = \mathbb{E}(e^{sX_1})\mathbb{E}(e^{sX_2}) \cdots \mathbb{E}(e^{sX_n}) = [M_X(s)]^n$$

So by law of total expectation,

$$\mathbb{E}(e^{sY}) = \sum_{n=1}^{\infty} \mathbb{E}(e^{sY} | N=n)\, p_N(n) = \sum_{n=1}^{\infty} [M_X(s)]^n p_N(n)$$

only difference

Recall that

$$M_N(s) = \mathbb{E}(e^{sN}) = \sum_{n=1}^{\infty} (e^s)^n p_N(n)$$

Thus we have a "recipe" for finding $M_Y(s) = \mathbb{E}(e^{sY})$:

- first find $M_N(s)$
- then plug $M_X(s)$ everywhere you see e^s in $M_N(s)$

Two cases where this yields "awesome" results.

Example

$N \sim$ geometric(p), $X \sim$ exponential(λ)

from last lecture

from last lecture

$$M_N(s) = \frac{pe^s}{1-(1-p)e^s}, \quad M_X(s) = \frac{\lambda}{\lambda - s}$$

Recipe $\Rightarrow$ $$M_Y(s) = \frac{\frac{p\lambda}{\lambda - s}}{1-(1-p)\frac{\lambda}{\lambda-s}} = \frac{p\lambda}{p\lambda - s} \Rightarrow Y \sim \text{exponential}(p\lambda)$$

Example

$N \sim$ geometric(p) ; $X \sim$ geometric(q)

$$M_N(s) = \frac{pe^s}{1-(1-p)e^s} \qquad M_X(s) = \frac{qe^s}{1-(1-q)e^s}$$

Recipe $\Rightarrow$ $$M_Y(s) = \frac{\frac{pqe^s}{1-(1-q)e^s}}{1-(1-p)\frac{qe^s}{1-(1-q)e^s}} = \frac{pqe^s}{1-(1-pq)e^s}$$

Next,

Limit Theorems

Focus on sums and averages of iid rvs

$$S_n = X_1 + X_2 + \dots + X_n \quad , X_k \text{ iid like } X$$

$$M_n = \frac{1}{n} S_n$$

Note: $\mathbb{E}(S_n) = n\mu \qquad \mathrm{Var}(S_n) = n\sigma^2$ ← both blow up as $n \to \infty$

constant as $n \to \infty$ goes to zero as $n \to \infty$

But: $\mathbb{E}(M_n) = \frac{1}{n}\mathbb{E}(S_n) = \mu$, $\quad \mathrm{Var}(M_n) = \frac{1}{n^2}\mathrm{Var}(S_n) = \frac{\sigma^2}{n}$

So pdf of M_n is always centered on μ, with spread $\to 0$ as $n \to \infty$

How exactly does this "asymptotic concentration" of the pmf or pdf of M_n around μ happen?

A useful fact (even though it's "weak" mathematically)

Chebyshev's Inequality: If X a rv w/ finite mean μ and finite variance $\mathrm{Var}(X)$, then

$$\mathbb{P}(\{|X-\mu| \geq c\}) \leq \frac{\mathrm{Var}(X)}{c^2} \qquad \forall c > 0$$

For large c, this is a "tail probability". As tail $c \to \infty$, goes to zero at least as fast as $\frac{1}{c^2}$.

Apply this to the M_n_converging_thing:

$$\mathrm{Var}(M_n) = \frac{\sigma^2}{n} \quad \forall\, n > 0$$

Thus $\forall\, \varepsilon > 0$, by Chebyshev we have

$$\mathbb{P}(\{|M_n - \mu| \geq \varepsilon\}) = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \to \infty} 0$$

A precise-ish statement of how concentration around μ happens

$$\forall\, \varepsilon > 0, \quad \lim_{n \to \infty} \mathbb{P}(\{|M_n - \mu| \geq \varepsilon\}) = 0$$

Known as the Weak Law of Large Numbers (WLLN)