

## Recall

- Conditional variance
- Law of total variance

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

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

Saw a couple of examples

i.e.

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

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

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. Instead use a table look-up.

### Example

Say  $X_1, X_2$  are Gaussian  $(\mu_1, \sigma_1^2)$  and  $(\mu_2, \sigma_2^2)$  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 ;  $Y = X_1 + X_2$

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

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

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

Here,

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

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

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

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  $n$ ;

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

where

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

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

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

Thus by law of iterated expectations

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

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

$$\text{Var}(Y|N=n) \stackrel{\text{by independence}}{=} n\text{Var}(X)$$

Thus

$$\text{Var}(Y|N) = N\text{Var}(X)$$

By law of total variance

$$\text{Var}(Y) = \mathbb{E}(\text{Var}(Y|N)) + \text{Var}(\mathbb{E}(Y|N))$$

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

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

Thus

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

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  $M_Y(s)$ .

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

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

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 some "awesome" results.

### Example

$N \sim \text{geometric}(p)$ ,  $X \sim \text{exponential}(\lambda)$

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

from last lecture

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

### Example

$N \sim \text{geometric}(p)$ ;  $X \sim \text{geometric}(q_f)$

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

$$M_X(s) = \frac{qe^s}{1 - (1-q_f)e^s}$$

$$\text{Recipe} \Rightarrow M_Y(s) = \frac{\frac{pfe^s}{1 - (1-q_f)e^s}}{1 - (1-p)\frac{qe^s}{1 - (1-q_f)e^s}} = \frac{pfe^s}{1 - (1-pq_f)e^s}$$

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:  $E(S_n) = n\mu$      $\text{Var}(S_n) = n\sigma^2$     both blow up as  $n \rightarrow \infty$

But:  $E(M_n) = \frac{1}{n} E(S_n) = \mu$ ,     $\text{Var}(M_n) = \frac{1}{n^2} \text{Var}(S_n) = \frac{\sigma^2}{n}$     \text{constant as } n \rightarrow \infty    \text{goes to zero as } n \rightarrow \infty

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

How exactly does this "asymptotic concentration" of the pdf or pdf of  $M_n$  around  $\mu$  happen?

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

Chebyshov's Inequality: If  $X$  a rv w/ finite mean  $\mu$  and finite variance  $\text{Var}(X)$ , then

$$P(|X - \mu| > c) \leq \frac{\text{Var}(X)}{c^2} \quad \forall c > 0$$

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

Apply this to the  $M_n$ -converging thing:

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

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

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

A precise-ish statement of how concentration around  $\mu$  happens

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

↑ Known as the Weak Law of Large Numbers (WLLN)