

① Probability Theory: Lecture #21

Nov. 19

Recall the definition of a moment for a discrete \mathbb{R} -valued r.v. X :

$$\begin{aligned} \text{the } p^{\text{th}} \text{ moment of } X &:= \mathbb{E}(X^p) \\ &= \sum_w [X(w)]^p \cdot P(w) \end{aligned}$$

$$\begin{aligned} \text{Note that the } 1^{\text{st}} \text{ and } 2^{\text{nd}} \text{ moments correspond to } \mathbb{E} \text{ and (assuming } \mathbb{E}=0) \\ \text{the variance: } \mu_1 &= \mathbb{E}(X^1) = \mathbb{E}(X) \\ \mu_2 &= \mathbb{E}(X^2) = \text{Var}(X) - \mathbb{E}(X)^2 \\ &= \text{Var}(X) \quad \text{when } \mathbb{E}(X)=0. \end{aligned}$$

Defn: Given a discrete r.v. $X: \Omega \rightarrow \mathbb{R}$, we define the moment-generating function of X to be

$$g(t) := \mathbb{E}(e^{tX}).$$

Why define this thing? Well, notice that the Taylor expansion

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!} \Rightarrow e^{tX} = \underbrace{\sum_{k=0}^{\infty} \frac{[tX]^k}{k!}}_{\text{r.v.}} = \underbrace{\sum_{k=0}^{\infty} \frac{t^k}{k!} X^k}_{\text{r.v.}}$$

$$\therefore \text{So, } g(t) = \mathbb{E}(e^{tX})$$

$$= \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{t^k}{k!} X^k\right)$$

$$= \sum_{k=0}^{\infty} \mathbb{E}\left(\frac{t^k}{k!} \underbrace{X^k}_{\text{random}}\right)$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{E}(X^k)$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \mu_k$$

In what sense does $g(t)$ generate moments?

Lemma : Let $g(t)$ be the MGF for the r.v. X . Then

$$\frac{d^k}{dt^k} [g(t)] \Big|_{t=0} = \mu_k .$$

Proof : Note that $\frac{d^k}{dt^k} [g(t)] = \sum_{k=0}^{\infty} \frac{d^k}{dt^k} [t^k] \cdot \frac{\mu_k}{k!}$

$$= k! \cdot \frac{\mu_k}{k!} + \text{Terms involving } t$$

$$= \mu_k + \text{Terms involving } t$$

$$\Rightarrow \frac{d^k}{dt^k} \left[g(t) \right] \Big|_{t=0} = \mu_k + 0$$

$$= \mu_k. \blacksquare$$

Ex (MGF of a binomial r.v.): Let $\bar{X} \sim \text{Bin}(n, p)$. Then

$$\begin{aligned} g(t) &= \mathbb{E}(e^{t\bar{X}}) = \binom{n}{m} \cdot p^m \cdot [1-p]^{n-m} \\ &= \sum_{m=0}^n e^{tm} \cdot P(\bar{X}=m) \\ &= \sum_{m=0}^n e^{tm} \cdot \binom{n}{m} \cdot p^m \cdot (1-p)^{n-m} \\ &= \sum_{m=0}^n [pe^t]^m [1-p]^{n-m} \binom{n}{m} \\ &\stackrel{\text{Binomial expansion}}{=} [pe^t + 1-p]^n. \end{aligned}$$

• Remarkably (or obviously, depending on your perspective), $g(t)$ uniquely determines \bar{X} !

Theorem (MGF determines r.v.): Let \bar{X} be a discrete r.v. with finite range $\{x_i\}_{i=1}^n$, distribution $p(x_i) = P(\bar{X}=x_i)$, and MGF $g(t)$. Then \bar{X} is the

only r.v. with MGF $g(t)$ with the same finite range.

Proof: Suppose Y has range $\{x_i\}_{i=1}^n$ and MGF $g(t)$. WLOG, assume $q(x_i) = P(Y=x_i) > 0$ for all x_i . Then we know

$$\begin{aligned} g(t) &= g(t) \\ \Rightarrow \cancel{\text{g}(t)} &\cancel{=} \cancel{\text{g}(t)} \end{aligned}$$

$$\mathbb{E}(e^{tX}) = \mathbb{E}(e^{tY})$$

$$\Rightarrow \sum_{i=1}^n e^{tx_i} p(x_i) = \sum_{i=1}^n e^{tx_i} q(x_i)$$

$$\Rightarrow \sum_{i=1}^n e^{tx_i} [p(x_i) - q(x_i)] = 0$$

Let $u = e^t$, so that the above may be written as

$$c_i = [p(x_i) - q(x_i)]$$

$$\sum_{i=1}^n u^{x_i} c_i = 0, \text{ for all } u$$

Suppose, wLOG that $x_1 < x_2 < x_3 < \dots < x_n$. We will show that each $c_i = 0$, which will give the result.

(5) Clearly, this result will follow if we can show $\{u^{x_i}\}_{i=1}^n$ are linearly independent as functions of u , for any choice of $x_1 < x_2 < \dots < x_n$.

This is established as follows:

Lemma: For any $x_1 < x_2 < \dots < x_n$, and any coefficients c_1, c_2, \dots, c_n ,

$$f(u) = \sum_{i=1}^n u^{x_i} \cdot c_i \quad \text{has at most } (n-1) \text{ roots, unless } c_1 = c_2 = \dots = c_n.$$

Proof: We proceed by induction. The case $n=1$ is clear, since $u^{x_1} = 0$ iff $u=0$.

Suppose the result holds for $(n-1)$. If $\sum_{i=1}^n u^{x_i} \cdot c_i$ violates the theorem by having $(n+1)$ or more roots. Then ~~the~~ the new function

$$F(u) = u^{-x_1} \cdot \sum_{i=1}^n u^{x_i} \cdot c_i$$

also has at least $(n+1)$ roots, since the u^{-x_1} only makes ~~worse~~ better.

Then $F(u) = 1 + \sum_{i=2}^n u^{y_i} \cdot c_i$ for different powers $y_i = x_i - x_1$.

$$\Rightarrow F'(u) = \sum_{i=2}^n u^{y_i-1} \cdot c_i$$

By induction, ~~F'~~ F' has at most $(n-1)$ roots. But if $F(u)$ has

N roots, $F'(u)$ has $N-1$ or more roots, by the MVT. Thus, (6)
 $F'(u)$ has at least $(n+1)-1 = n$ roots. Contradiction. \blacksquare Lemma
 \square Theorem