Lecture 21

Generating functions

Moment generating functions

Moments and the MGF

Sums of independent variables

Characteristic functions

Probability generating functions

Generating functions

Probability generating function

$$g_X(s) \;=\; E\left\{s^X\right\} \;=\; \sum_x s^x \; p_X(x)$$

Moment generating function

$$M_X(t) \;=\; E\left\{e^{tX}\right\} \;=\; \sum_x e^{tx} \; p_X(x)$$

Characteristic function

$$\varphi_X(\theta) \;=\; E\left\{e^{i\theta X}\right\} \;=\; \sum_x e^{i\theta x} \; p_X(x)$$

where $i = \sqrt{-1}$.

Moment generating functions

$$M_X(t) \;=\; E\left\{e^{tX}\right\} \;=\; \sum_x e^{tx}\, p_X(x)$$

X Bernoulli, parameter p

$$M_X(t) \;=\; e^t p + e^0 q \;=\; pe^t + q$$

$X \sim \text{Bin}(n, p)$

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k q^{n-k} \\ &= (pe^t + q)^n \end{aligned}$$

$X \sim \text{Poiss}(\lambda)$

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{\infty} e^{tk} \left(\frac{e^{-\lambda}\lambda^k}{k!} \right) \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} \\ &= e^{-\lambda} e^{\lambda e^t} = e^{-\lambda(1-e^t)} \end{aligned}$$

$X \sim \text{Geom}(p)$

$$\begin{aligned} M_X(t) = \sum_{k=1}^{\infty} e^{tk}\, pq^{k-1} &= pe^t \sum_{k=1}^{\infty} (qe^t)^{k-1} \\ &= pe^t \sum_{k=0}^{\infty} (qe^t)^k = \frac{pe^t}{1-qe^t} \end{aligned}$$

Finding the moments from the MGF

$$\begin{aligned}
\frac{d}{dt} M_X(t) &= \frac{d}{dt} \sum_x e^{tx}\ p_X(x) \\
&= \sum_x \frac{d}{dt}\left(e^{tx}\right)\ p_X(x) \\
&= \sum_x x\ e^{tx}\ p_X(x)
\end{aligned}$$

So

$$\frac{d}{dt} M_X(t)\Big|_{t=0} = \sum_x x\ p_X(x) = E\{X\}$$

Similarly

$$\frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = \sum_x x^2\ p_X(x) = E\{X^2\}$$

$$\frac{d}{dt}M_X(t)\Big|_{t=0} = E\{X\}, \quad \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} = E\{X^2\}$$

Binomial, $M_X(t) = (pe^t + q)^n$

$$M_X'(t) = npe^t(pe^t + q)^{n-1} \quad \Rightarrow \quad E\{X\} = M_X'(0) = np$$

$$M_X''(t) = npe^t(pe^t + q)^{n-1} + n(n-1)(pe^t)^2(pe^t + q)^{n-2}$$

$$\Rightarrow E\{X^2\} = M_X''(0) = (np)^2 + npq$$

Poisson, $M_X(t) = e^{-\lambda(1-e^t)}$

$$M_X'(t) = \lambda e^{-t}e^{-\lambda(1-e^t)} \quad \Rightarrow \quad E\{X\} = M_X'(0) = \lambda$$

$$M_X''(t) = \lambda e^{-t}e^{-\lambda(1-e^t)} + \lambda^2 e^{-2t}e^{-\lambda(1-e^t)}$$

$$\Rightarrow E\{X^2\} = M_X''(0) = \lambda + \lambda^2$$

Moments and the MGF

- FACT: If $M_X(t)$ exists for all $t \in [-\epsilon, +\epsilon]$, then ALL the moments $E\{X^k\}$ exist and are finite.

$$
\begin{aligned}
E\{X^n\} &= \sum_{k=-\infty}^{\infty} k^n \; p_X(k) \\
&= \sum_{-\infty}^{-1} k^n \; p_X(k) \; + \; \sum_{1}^{\infty} k^n \; p_X(k) \\
&= I \; + \; II
\end{aligned}
$$

$$
\begin{aligned}
II &\leq \sum_{k:k^n \geq \exp(\epsilon k)} k^n p_X(k) \; + \sum_{k:k^n < \exp(\epsilon k)} e^{\epsilon k} p_X(k) \\
&\leq C(n, \epsilon) \; + \; M_X(\epsilon) \; < \; \infty
\end{aligned}
$$

The other term is similar, but uses $-\epsilon$.

Characteristic functions

$$\varphi_X(\theta) \;=\; E\{e^{i\theta X}\} \;=\; E\{e^{(i\theta)X}\} \;=\; M_X(i\theta)$$

- MGF's do not always exist, since moments do not.
- Characteristic functions always exist.

$$\begin{aligned}
\left|E\{e^{i\theta X}\}\right| \;=\; \left|\sum_x e^{i\theta x}\; p_X(x)\right| \;&\leq\; \sum_x \left|e^{i\theta x}\right| p_X(x) \\
&=\; \sum_x p_X(x) \\
&=\; 1
\end{aligned}$$

- Derivatives of characteristic functions evaluated at $\theta = 0$ give moments.

$$\frac{d}{d\theta}\varphi_X(\theta)\Big|_{\theta=0} = iE\{X\}, \quad \frac{d^2}{d\theta^2}\varphi_X(\theta)\Big|_{\theta=0} = -E\{X^2\}, \quad \text{etc}$$

Uniqueness of generating functions

- Each probability function generates only one MGF or CF. (Obvious)

- Each characteristic function comes from only one probability function. (Deep result in complex analysis)

- Each moment generating function (along with its region of convergence) comes from only one probability function. (Deep result in real analysis)

- Using the definitions
  $$M_X(t) = E\left\{e^{tX}\right\}, \ \varphi_X(\theta) = E\left\{e^{i\theta X}\right\},$$
  the same uniqueness will hold for "continuous random variables".

MGF's of sums

- $X_1, \ldots, X_n$ independent with MGF's $M_{X_1}(t), \ldots, M_{X_n}(t)$.
- $S_n = X_1 + \ldots + X_n$

$$\boxed{M_{S_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t)}$$

$$\begin{aligned}
M_{S_n}(t) &= E\{e^{tS_n}\} \\
&= E\{e^{t(X_1+\ldots+X_n)}\} \\
&= E\{e^{tX_1} \cdots e^{tX_n}\} \\
&= E\{e^{tX_1}\} \cdots E\{e^{tX_n}\} \\
&= M_{X_1}(t) \cdots M_{X_n}(t)
\end{aligned}$$

Applications

- $X \sim \text{Bin}(n_1, p)$, independent of $Y \sim \text{Bin}(n_2, p)$,
  $\Rightarrow \ (X + Y) \sim \text{Bin}(n_1 + n_2, p)$.

$$\begin{aligned} M_{X+Y}(t) \ = \ M_X(t) M_Y(t) \ &= \ (pe^t + q)^{n_1} \ (pe^t + q)^{n_2} \\ &= \ (pe^t + q)^{n_1 + n_2} \end{aligned}$$

- $X \sim \text{Poiss}(\lambda_1)$, independent of $Y \sim \text{Poiss}(\lambda_2)$,
  $\Rightarrow \ (X + Y) \sim \text{Poiss}(\lambda_1 + \lambda_2)$.

$$\begin{aligned} M_{X+Y}(t) \ = \ M_X(t) M_Y(t) \ &= \ e^{-\lambda_1 (1 - e^t)} \ e^{-\lambda_2 (1 - e^t)} \\ &= \ e^{-(\lambda_1 + \lambda_2)(1 - e^t)} \end{aligned}$$

Probability generating functions

$$g_X(s) \;=\; E\{s^X\} \;=\; E\{e^{\ln(s)X}\} \;=\; M_X(\ln(s))$$

- Assume that X takes values only in $\{0, 1, 2, \ldots\}$.
- $g_X(s)$ exists for all $s \in [-1, 1]$.

$$|g_X(s)| \;=\; \left|\sum_{k\geq 0} s^k p_X(k)\right| \;\leq\; \sum_{k\geq 0} |s|^k p_X(k) \;\leq\; \sum_{k\geq 0} p_X(k) \;=\; 1$$

- Derivatives of $g_X(s)$ at $s = 1$ give moments.
- $X_1, \ldots, X_n$ independent $\Rightarrow$
  $g_{X_1+\cdots X_n}(s) = g_{X_1}(s) \cdots g_{X_n}(s)$, etc, etc.

- Derivatives of $g_X(s)$ at $s = 0$ give probabilities!

$$\begin{aligned}
\frac{d}{ds}\, g_X(s) &= \frac{d}{ds} \sum_{k \geq 0} s^k \; p_X(k) \\
&= \sum_{k \geq 0} \frac{d}{ds} \left(s^k\right) \; p_X(k) \\
&= \sum_{k \geq 1} k \; s^{k-1} \; p_X(k) \\
\Rightarrow \quad \frac{d}{ds}\, g_X(s)\Big|_{s=0} &= p_X(1)
\end{aligned}$$

And, continuing,

$$\frac{d^n}{ds^n}\, g_X(s)\Big|_{s=0} = n! \; p_X(n)$$

- Good news: Computers can differentiate