2 Random variables and distributions (cont'd)

2.6 Indicator (cont'd)

$$I(w) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \not\in A \end{cases}$$

$$P(I_A) = P(A)$$

Example 3

Let X be a r.v. taking values in non-negative integers, then

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} P(X > n)$$

Proof:

Note that $X = \sum_{n=0}^{\infty} I_n$ where $I_n = I_{x>n}$. ($x > n$ is an event)

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(\sum_{n=0}^{\infty} I_n) \\ &= \sum_{n=0}^{\infty} \mathbb{E}(I_n) \\ &= \sum_{n=0}^{\infty} P(X > n) \end{aligned}$$

In particular, let $X \sim Geo(p)$. As we have seen, $P(X > n) = (1-p)^n \Rightarrow$

$$\begin{aligned} \mathbb{E}(X) &= \sum_{n=0}^{\infty} P(X > n) \\ &= \sum_{n=0}^{\infty} (1-p)^n \\ &= \frac{1}{1-(1-p)} = \frac{1}{p} \end{aligned}$$

2.7 Moment generating function

Definition: Let X be a r.v. Then the function $M(t) = \mathbb{E}(e^{tx})$ is called the *moment generating function(mgf)* of X, if the expectation exists for all $t \in (-h, h)$ for some $h > 0$.

Remark: The mgf is not always well-defined. It is important to check the existence of the expectation.

Properties of mgf

1. Moment Generating Function generates moments
    - **Theorem**:

- $M(0) = 1$
- $M^{(k)}(0) = \mathbb{E}(X^k), k = 1, 2, \dots$ ($M^{(k)} = \frac{d^k}{dt^k} M(t)|_{t=0}$)
  - **Proof**:

$$\begin{aligned} M(0) &= \mathbb{E}(e^{0 \cdot X}) = \mathbb{E}(1) = 1 \\ M^{(k)}(0) &= \frac{d^k}{dt^k} \mathbb{E}(e^{t \cdot X})|_{t=0} \\ &= \mathbb{E}(\frac{d^k}{dt^k} e^{tX}|_{t=0}) \\ &= \mathbb{E}(X^k) \end{aligned}$$

- As a result, we have: $M(t) = \sum_{k=0}^{\infty} \frac{M^{(k)}(0)}{k!} t^k = \sum_{k=0}^{\infty} \frac{E * X^k}{k!} t^k$ (a method to get moment of a r.v)

2. $X \perp\!\!\!\perp Y$, with mgf's M_x, M_y. Let M_{X+Y} be the mgf of $X + Y$. then

$$M_{X+Y}(t) = M_X(t) M_Y(y)$$

- **Proof**:

$$\begin{aligned} M_{X+Y}(t) &= \mathbb{E}(e^{t(X+Y)}) \\ &= \mathbb{E}(e^{tx} e^{ty}) \\ &= \mathbb{E}(e^{tx}) \mathbb{E}(e^{ty}) \\ &= M_X(y) M_Y(t) \end{aligned}$$

3. The mgf completely determines the distribution of a r.v.
   - $M_X(t) = M_Y(t)$ for all $t \in (-h, h)$ for some $h > 0$, then $X \overset{d}{=} Y$. ($\overset{d}{=}$: have the smae distribution)
   - Example: Let $X \sim Poi(\lambda_1)$, $Y \sim Poi(\lambda_2)$. $X \perp\!\!\!\perp Y$. Find the distribution of $X + Y$
     - First, derive the mgf of a Poisson distribution.

$$\begin{aligned} M_X(t) &= \mathbb{E}(e^{tX}) \\ &= \sum_{n=0}^{\infty} e^{tn} \cdot P(X = n) \\ &= \sum_{n=0}^{\infty} e^{tn} \cdot \frac{\lambda_1^n}{n!} e^{-\lambda_1} \\ &= \sum_{n=0}^{\infty} \frac{(e^t \cdot \lambda_1)^n}{n!} \cdot e^{-\lambda_1} \end{aligned}$$

$$\text{we know that } \sum_{n=0}^{\infty} \frac{(e^t \lambda_1)^n}{n!} = e^{e^t \cdot \lambda_1}. (\text{Since } \frac{(e^t \lambda_1^n)}{n!} e^{-e^t \lambda_1} \text{ is the pmf of } Poi(e^t \lambda_1))$$

$$\Rightarrow M_X(t) = e^{e^t \lambda_1} e^{-\lambda_1} = e^{\lambda_1 (e^t - 1)}, t \in \mathbb{R}. (e^{\lambda_1 (e^t - 1)} \text{ is mgf of } Poi(\lambda_1))$$

$$\text{Similarly, } M_Y(t) = e^{\lambda_2 (e^t - 1)}.$$

We know that

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

This is the mgf of $Poi(\lambda_1 + \lambda_2)$!

Since the mgf uniquely determines the distribution $X + Y \sim Poi(\lambda_1 + \lambda_2)$

In general, if $X_1, X_2, \ldots, X_n$ independent, $X_i \sim Poi(\lambda_i)$, then $\sum X_i \sim Poi(\sum \lambda_i$

2.7.2 Joint mgf

Definition: Let X, Y be r.v's. Then $M(t_1, t_2) := \mathbb{E}(e^{t_1 X + t_2 Y}$ is called the joint mgf of X and Y, if the expectation exists for all $t_1 \in (-h_1, h_1)$, $t_2 \in (-h_2, h_2)$ for some $h_1, h_2 > 0$.

More generally, we can define $M(t_1, \ldots, t_n) = \mathbb{E}(exp(\sum_{i=1}^n t_i X_i))$ for r.v's $X_1, \cdots, X_n$, if the expectation exists for $\{(t_1, \cdots, t_n) : t_i \in (-h_i, h_i), i = 1, \cdots, n\}$ for some $\{h_i > 0\}, i = 1, \cdots, n$

Properties of the joint mgf

1. $$\begin{aligned} M_X(t) &= \mathbb{E}(e^{tX}) \\ &= \mathbb{E}(e^{tX + oY}) \\ &= M(t, o) \\ M_Y(t) &= M(o, t) \end{aligned}$$

2. $$\frac{\partial^{m+n}}{\partial t_1^m \partial t_2^n} M(t_1, t_2)|_{(0,0)} = \mathbb{E}(X^m Y^n \\ \text{the proof is similar to the single r.v. case}$$

3. If $X \perp\!\!\!\perp Y$, then $M(t_1, t_2) = M_X(t_1) M_Y(t_2)$
   - **Proof**:

   $$\begin{aligned} M(t_1, t_2) &= \mathbb{E}(e^{t_1 X + t_2 Y}) \\ (X \perp\!\!\!\perp Y) &= \mathbb{E}(e^{t_1 X} e^{t_2 Y}) \\ &= \mathbb{E}(e^{t_1 X}) \cdot \mathbb{E}(e^{t_2 Y}) \\ &= M_X(t_1) \cdot M_Y(t_2) \end{aligned}$$

   - **Remark**: Don't confuse this with the result $X \perp\!\!\!\perp Y \Rightarrow M_{X+Y}(t) = M_X(t) M_Y(t)$.
     - $M_{X+Y}(t) \rightarrow$ mgf of $X + Y$; single argument function t
     - $M(t_1, t_2) \rightarrow$ joint mgf of (X, Y); two arguments t_1, t_2