**Moment-Generating Functions** MGF → r.v. **Notes 08**

**Associated Reading: Wackerly 7, Chapter 3, Section 9, Chapter 4, Section 9, and Chapter 6, Section 5**

In this notes set, we introduce an alternative means of specifying a probability distribution: the *moment-generating function*, or mgf.

Let's first put the idea of alternative specifications into context. *They exist as tools that allow one to derive analytical results that may not be as easily derived using either a distribution's pmf/pdf or its cdf.* The mgf is not the only alternative specification; there are also, e.g.,

- the *probability-generating function*, for discrete r.v.'s (see Wackerly 7, Chapter 3, Section 10);
- the *characteristic function*, the inverse Fourier transform of a pmf/pdf; and
- the *cumulant-generating function*. $K(t) = \log \mathbb{E}[e^{tY}]$.

Y ~ r.v of interest.

$\psi_Y(t) = \mathbb{E}[e^{itY}]$

i ⇒ imaginary number, $i^2 = -1$

To begin our coverage of mgfs, we introduce the concept of *moments*.

- The $k^{\text{th}}$ moment of a r.v. $Y$ about the origin is:

$$\mu'_k = \mathbb{E}[Y^k] = \begin{cases} \sum_{D_y} y^k p(y) \\ \int_{D_y} y^k f(y)\, dy \end{cases}$$

k=1 first moment. $\mu'_1 = \mathbb{E}[Y] = \mu$. mean of dist.

- The $k^{\text{th}}$ moment of a r.v. $Y$ about the mean, or its $k^{\text{th}}$ central moment, is:

$$\mu_k = \mathbb{E}[(Y-\mu)^k] = \begin{cases} \sum_{D_y} (y-\mu)^k p(y) \\ \int_{D_y} (y-\mu)^k f(y)\, dy. \end{cases}$$

k=1 $\mu_1 = 0$

k=2 $\mu_2 = V[Y] = \mathbb{E}(Y-\mu)^2$

The moments of a given distribution are unique, i.e., if $\mu'_{X,i} = \mu'_{Y,i}\ \forall\, i \in \mathbb{Z}^+$, then the probability distributions for the r.v.'s $X$ and $Y$ are *identical*.

The moments of a given distribution can be encapsulated in an mgf:

moment-generating function.

$$\boxed{m_Y(t)} = \mathbb{E}[e^{tY}]. \qquad e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$$

$$= \mathbb{E}\left[1 + (tY) + \frac{(tY)^2}{2} + \frac{(tY)^3}{3!} + \cdots\right]$$

$$= 1 + t\cdot\mathbb{E}[Y] + t^2\cdot\frac{\mathbb{E}Y^2}{2!} + t^3\cdot\frac{\mathbb{E}Y^3}{3!} + \cdots \infty = \sum_{k=0}^{\infty} \frac{t^k \mu'_k}{k!}$$

$\mu'_1$ $\quad \mu'_2/2!$ $\quad \mu'_3$

Note that an mgf only exists for a particular distribution if there is a constant $b$ such that $m(t)$ is finite for $|t| < b$. An example of a distribution for which the mgf does not exist is the Cauchy distribution.

↳ doesn't have expectation EX

$m_X(t) = m_Y(t) \Rightarrow \mu'_{Xk} = \mu'_{Yk} \Rightarrow X, Y$ same

$$\sum_{k=0}^{\infty} \frac{t^k \cdot \mu_{Xk}'}{k!} = \sum_{k=0}^{\infty} \frac{t^k \cdot \mu_{Yk}'}{k!}$$

dist.



By Theorem 3.12, if we are given a distribution's mgf, we can use it to derive any moment about the origin, $\mu_i'$, of that distribution:

Claim: $\boxed{\left.\frac{d^k m_Y(t)}{(dt)^k}\right|_{t=0} = \mu_k'}$

$$\left.\frac{d^k m_Y(t)}{(dt)^k}\right|_{t=0} = \frac{d^k}{(dt)^k}\left(1 + \mu_1' t + \frac{\mu_2'}{2!}t^2 + \frac{\mu_3'}{3!}t^3 + \cdots\right)$$

$$= 0 + \frac{k!\,\mu_k'}{k!} + 0$$

$$= \mu_k'$$

Basic property: $\frac{d^k(t^k)}{(dt)^k} = k!$

$\frac{d^k(t^{k-1})}{(dt)^k} = 0$

$\frac{d^k(t^{k+1})}{(dt)^k} = (k+1)\cdot k \cdot (k-1) \cdots 2t$

→ **EXAMPLE.** Wackerly 7, Exercises 3.145 and 3.146

**3.145** If $Y$ has a binomial distribution with $n$ trials and probability of success $p$, show that the moment-generating function for $Y$ is

$$m(t) = (pe^t + q)^n, \qquad \text{where } q = 1 - p.$$

**3.146** Differentiate the moment-generating function in Exercise 3.145 to find $E(Y)$ and $E(Y^2)$. Then find $V(Y)$.

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

$$m_Y(t) = E[e^{tY}]$$

$$= \sum_{y=0}^{n} e^{ty} p(y) \nearrow \binom{n}{y} p^y (1-p)^{n-y}, \quad 1-p = q$$

$$= \sum_{y=0}^{n} e^{ty} \binom{n}{y} p^y q^{n-y}$$

$$= \sum_{y=0}^{n} \binom{n}{y} (e^t \cdot p)^y \cdot q^{n-y}$$

$$= (e^t \cdot p + q)^n$$

recall: Binomial theorem, $(a+b)^n = \sum_{y=0}^{n} \binom{n}{y} a^y \cdot b^{n-y}$

let $a = e^t \cdot p$

$b = q$

$$\boxed{E[Y]} = \mu_1' = \left.\frac{d\,m_Y(t)}{dt}\right|_{t=0}$$

$$= \left. n(e^t \cdot p + q)^{n-1} \cdot p \cdot e^t \right|_{t=0}$$

$$= n \cdot (p+q)^{n-1} \cdot p, \quad p+q = 1$$

$$= np$$

$$V[Y] = EY^2 - (EY)^2$$

$$\boxed{E[Y^2]} = \mu_2' = \left.\frac{d^2(m_Y(t))}{(dt)^2}\right|_{t=0}$$

$$= \left.\left(n(e^t \cdot p + q)^{n-1} \cdot p \cdot e^t\right)'\right|_{t=0}$$

$$= n^2p^2 - np^2 + np$$

$$\boxed{V[Y] = np(1-p)}$$

→ **EXAMPLE.** Wackerly 7, Exercise 4.140

**4.140** Identify the distributions of the random variables with the following moment-generating functions:

a $m(t) = (1 - 4t)^{-2}$.

b $m(t) = 1/(1 - 3.2t)$.

c $m(t) = e^{-5t+6t^2}$.

mgf → r.v.

Just match with the table!

→ **EXAMPLE.** What is the mgf for the gamma distribution?

$Y \sim Gamma(\alpha, \beta)$

$\boxed{m_Y(t)} = \mathbb{E}[e^{tY}]$

$= \int_0^\infty e^{ty} f(y)\, dy$

$= \int_0^\infty e^{ty} \cdot \frac{y^{\alpha-1} \cdot e^{-\frac{y}{\beta}}}{\beta^\alpha \cdot \Gamma(\alpha)}\, dy.$

$(t - \frac{1}{\beta}) y = - \frac{y}{\frac{1}{\frac{1}{\beta} - t}}$

$\alpha > 0 \quad \beta' > 0 \quad t < \frac{1}{\beta}$

$Gamma(\alpha, \beta' = \frac{1}{\frac{1}{\beta} - t})$

$= \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} \boxed{\int_0^\infty y^{\alpha-1} e^{(ty - \frac{y}{\beta})} dy.} = \frac{1}{\beta^\alpha \cancel{\Gamma(\alpha)}} \cdot (\beta')^\alpha \cdot \cancel{\Gamma(\alpha)}$

$e^{-\frac{y}{\beta'}}$

$= (\frac{\beta'}{\beta})^\alpha = \left(\frac{1}{(\frac{1}{\beta} - t)\beta}\right)^\alpha = \boxed{(1 - \beta t)^{-\alpha}}$

$t < \frac{1}{\beta}.$

There is one last piece of unfinished business before we transition to the use of mgfs to find distributions of functions of random variables. The Law of the Unconscious Statistician extends to the computation of mgf's:

$U = g(Y).$

suppose $Y$ is continuous r.v.

$m_U(t) = \mathbb{E}[e^{tg(Y)}] = \int_{D_y} e^{tg(y)} f(y)\, dy$ (general fution of $g(Y)$).

particularly if $g(Y) = aY + b$. ($a, b$ determinist constants).

$m_{g(Y)}(t) = \mathbb{E}[e^{t(aY+b)}] = \mathbb{E}[e^{atY} \cdot e^{t \cdot b}] = e^{bt} \mathbb{E}[e^{(at) \cdot Y}]$

$m_Y(at)$.

$$\boxed{m_{g(Y)}(t) = e^{bt} m_Y(at).}$$

eg. $Y \sim Exp(4)$

$U = 3Y + 1.$

$\boxed{m_U(t)} = e^t m_Y(3t)$

$= e^t \cdot (1 - 4 \cdot 3t)^{-1} = \boxed{\frac{e^t}{1 - 12t}}$

If $\boxed{S_n = \sum_{i=1}^n a_i Y_i}$, where the $Y_i$ are independent but not necessarily identically distributed r.v.'s, then:

method of dist. $P(S_n \le u) = P(\sum_{i=1}^n a_i Y_i \le u) = \int\int\int\int \cdots \int_{\sum_{i=1}^n a_i y_i \le u} f(y_1 \cdots y_n)\, dy_1 dy_2 \cdots dy_n$

cdf for $S_n$

join density of $Y_1 Y_2 \cdots Y_n$

method of mgf: $m_{S_n}(t) = \mathbb{E}[e^{t \cdot \sum_{i=1}^n a_i Y_i}] = \mathbb{E}[e^{ta_1 Y_1} \cdot e^{ta_2 Y_2} \cdots e^{ta_n Y_n}]$

$Y_1 \cdots Y_n$ are independent $= \prod_{i=1}^n \mathbb{E}[e^{ta_i Y_i}]$

$m_{Y_i}(a_i t)$.

$$\boxed{m_{S_n}(t) = \prod_{i=1}^n m_{Y_i}(a_i t)}$$

$Y_1\ Y_2$ independent

$\Rightarrow \mathbb{E}[f(Y_1)\, g(Y_2)] = \mathbb{E} f(Y_1) \cdot \mathbb{E} g(Y_2)$ (*)



→ **EXAMPLE.** What is the distribution of $Y_1 - Y_2$ when $Y_1 \sim \text{Pois}(\lambda_1)$ and $Y_2 \sim \text{Pois}(\lambda_2)$, and $Y_1$ and $Y_2$ are independent r.v.'s? Also, using the mgf, show that the expected value of $Y_1 - Y_2$ is $\lambda_1 - \lambda_2$.

$Y_1 \sim Pois(\lambda_1)$

$Y_2 \sim Pois(\lambda_2)$

$S_n = Y_1 - Y_2$.

$= a_1 Y_1 + a_2 Y_2$

$a_1 = 1 \quad a_2 = -1$.

$$\boxed{m_{S_n}(t)} = m_{Y_1}(t) \cdot m_{Y_2}(-t). \quad \text{by (*)}$$

$$= e^{\lambda_1(e^t - 1)} \cdot e^{\lambda_2(e^{-t}-1)}. \quad \text{by table.}$$

$$= \boxed{e^{\lambda_1 e^t + \lambda_2 e^{-t} - (\lambda_1 + \lambda_2)}}.$$

mgf for skellam distribution.

$$\mathbb{E}[S_n] = \frac{d\, m_{S_n}(t)}{dt}\Big|_{t=0}.$$

$$= e^{\lambda_1 e^t + \lambda_2 e^{-t} - (\lambda_1+\lambda_2)} \cdot (\lambda_1 e^t - \lambda_2 e^{-t})\Big|_{t=0}$$

$$= e^{(\lambda_1+\lambda_2) - (\lambda_1+\lambda_2)} (\lambda_1 - \lambda_2)$$

$$= \lambda_1 - \lambda_2 \quad ✓$$

$\mathbb{E}[Y_1 - Y_2]$

$= \mathbb{E}[Y_1] - \mathbb{E}[Y_2]$

$= \lambda_1 - \lambda_2$.

→ **EXAMPLE.** The sum of independent normally distributed r.v.'s follows what distribution?

$Y_i \sim N(\mu_i, \sigma_i^2)$.

$S_n = \sum_{i=1}^n Y_i$

$= \sum_{i=1}^n a_i Y_i$

$a_i = 1$.

$$\boxed{m_{S_n}(t)} = \prod_{i=1}^n m_{Y_i}(t).$$

$$= \prod_{i=1}^n \exp\left(\mu_i t + \tfrac{1}{2}\sigma_i^2 t^2\right).$$

$$= \exp\left[\sum_{i=1}^n \left(\mu_i t + \tfrac{1}{2}\sigma_i^2 t^2\right)\right]$$

$$= \exp\Big[\underbrace{\Big(\sum_{i=1}^n \mu_i\Big)}_{\mu'} t + \tfrac{1}{2}\underbrace{\Big(\sum_{i=1}^n \sigma_i^2\Big)}_{\sigma'^2} t^2\Big].$$

$$m_{S_n}(t) = \exp\left(\mu' t + \tfrac{1}{2}\sigma'^2 t^2\right)$$

$$\sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right).$$

⚠ fact: the sum of independent normal r.v's is itself a normal r.v.

⚠ fact: the sum of $n$ independent $\chi^2(1)$. $\Rightarrow \chi^2(n)$.

→ **EXAMPLE.** The sum of the squares of independent standard normal-distributed r.v.'s follows what distribution?

$Z_i \sim N(0,1)$.

$S_n = \sum_{i=1}^n Z_i^2$

$= \sum_{i=1}^n a_i Z_i^2$

for $a_i = 1$.

in previous notes, we demonstrated that

$Z \sim N(0,1) \quad Z^2 \sim \chi^2(1)$. (chi-square).

Gamma$(\frac{1}{2}, 2)$.

$m_{S_n}(t) = \prod_{i=1}^n m_{Z_i^2}(t)$

$= \prod_{i=1}^n \cdot (1-2t)^{-\frac{1}{2}}$

$= (1-2t)^{-\frac{n}{2}}$. $\Rightarrow S_n \sim \chi^2(n)$

→ **EXAMPLE.** Wackerly 7, Exercise 6.43

**6.43** Refer to Exercise 6.41. Let $Y_1, Y_2, \ldots, Y_n$ be independent, normal random variables, each with mean $\mu$ and variance $\sigma^2$. $N(\mu, \sigma^2)$.

**a** Find the density function of $\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$.

**b** If $\sigma^2 = 16$ and $n = 25$, what is the probability that the sample mean, $\overline{Y}$, takes on a value that is within one unit of the population mean, $\mu$? That is, find $P(|\overline{Y} - \mu| \le 1)$.

**c** If $\sigma^2 = 16$, find $P(|\overline{Y} - \mu| \le 1)$ if $n = 36$, $n = 64$, and $n = 81$. Interpret the results of your calculations.

a). $Y_i \sim N(\mu, \sigma^2)$

$\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$

$= \sum_{i=1}^n a_i Y_i$

$\Rightarrow a_i = \frac{1}{n}$

$m_{\overline{Y}}(t) = \prod_{i=1}^n m_{Y_i}(\frac{1}{n}t)$

$= \prod_{i=1}^n \exp(\mu \cdot \frac{1}{n}t + \frac{1}{2}\sigma^2 \cdot (\frac{1}{n}t)^2)$.

$= \exp\left( \sum_{i=1}^n \mu \frac{1}{n}t + \sum_{i=1}^n \left(\frac{1}{2}\sigma^2\right) \frac{1}{n^2}t^2 \right)$.

$= \boxed{\exp\left( \mu \cdot t + \frac{1}{2} \cdot \frac{\sigma^2}{n} t^2 \right)}$

$\overline{Y} \sim N(\mu, \frac{\sigma^2}{n})$.

cdf $\Phi(\cdot)$

$N(0,1)$.

$\Uparrow$

b) $\sigma^2 = 16$, $n = 25$.

$\boxed{P(|\overline{Y} - \mu| \le 1)} = P(-1 \le \overline{Y} - \mu \le 1) = P\left( \frac{-1}{\sqrt{\frac{\sigma^2}{n}}} \le \boxed{\frac{\overline{Y}-\mu}{\sqrt{\frac{\sigma^2}{n}}}} \le \frac{1}{\sqrt{\frac{\sigma^2}{n}}} \right)$

$\Phi(x) := P(Z \le x)$.

$= \Phi\left(\frac{1}{\sqrt{\frac{\sigma^2}{n}}}\right) - \Phi\left(-\frac{1}{\sqrt{\frac{\sigma^2}{n}}}\right)$

recall $\Phi(x) + \Phi(-x) = 1$

$\boxed{= 2\Phi\left(\frac{\sqrt{n}}{\sigma}\right) - 1}$ ✓

plug in $\sigma^2 = 16$ $n = 25$

$\Rightarrow P(|\overline{Y} - \mu| \le 1)$

$= 0.7887$.

$\overline{Y} \in [\mu - 1, \mu + 1]$

c).

| $n$ | $P(\overline{Y} \in [\mu-1, \mu+1]) = 2\Phi(\frac{\sqrt{n}}{\sigma}) - 1$ |
| --- | --- |
| 36 | 0.866 |
| 64 | 0.955 |
| 81 | 0.976 |

$\Rightarrow$ as $n \uparrow$ estimate of $\mu$ gets more precise

→ **REVIEW:** in what sort of problems can mgf's appear?

① Given a distribution, compute the mgf.

$$\rightarrow \quad m_Y(t) = \mathbb{E}[e^{tY}]$$

② Given a mgf, compute $\mathbb{E}[Y]$, $V[Y]$, $\mathbb{E}[Y^k]$.

$$\rightarrow \quad \left.\frac{d^k m_Y(t)}{(dt)^k}\right|_{t=0} \Rightarrow \underline{\mu_k'}$$

③ Given $\underbrace{U = \sum_{i=1}^{n} a_i Y_i}_{Y_i \text{ are independent.}}$, compute its dist

**Continuous Distributions**

| Distribution | Probability Function | Mean | Variance | Moment-Generating Function |
|---|---|---|---|---|
| Uniform | $f(y) = \frac{1}{\theta_2 - \theta_1}; \theta_1 \le y \le \theta_2$ | $\frac{\theta_1 + \theta_2}{2}$ | $\frac{(\theta_2 - \theta_1)^2}{12}$ | $\frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$ |
| Normal | $f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(y-\mu)^2\right]$ $-\infty < y < +\infty$ | $\mu$ | $\sigma^2$ | $\exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)$ |
| Exponential | $f(y) = \frac{1}{\beta} e^{-y/\beta}; \quad \beta > 0$ $0 < y < \infty$ | $\beta$ | $\beta^2$ | $(1 - \beta t)^{-1}$ |
| Gamma | $f(y) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right] y^{\alpha-1} e^{-y/\beta};$ $0 < y < \infty$ | $\alpha\beta$ | $\alpha\beta^2$ | $(1 - \beta t)^{-\alpha}$ |
| Chi-square | $f(y) = \frac{(y)^{(v/2)-1} e^{-y/2}}{2^{v/2}\Gamma(v/2)};$ $y^2 > 0$ | $v$ | $2v$ | $(1-2t)^{-v/2}$ |
| Beta | $f(y) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] y^{\alpha-1}(1-y)^{\beta-1};$ | $\frac{\alpha}{\alpha+\beta}$ | $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ | does not exist in closed form |

MGF = $(1-4t)^{-2}$

↓ rv.

Gamma ($\alpha = 2$, $\beta = 4$).

**Discrete Distributions**

| Distribution | Probability Function | Mean | Variance | Moment-Generating Function |
|---|---|---|---|---|
| Binomial | $p(y) = \binom{n}{y} p^y (1-p)^{n-y};$ $y = 0, 1, \ldots, n$ | $np$ | $np(1-p)$ | $[pe^t + (1-p)]^n$ |
| Geometric | $p(y) = p(1-p)^{y-1};$ $y = 1, 2, \ldots$ | $\frac{1}{p}$ | $\frac{1-p}{p^2}$ | $\frac{pe^t}{1-(1-p)e^t}$ |
| Hypergeometric | $p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}};$ $y = 0, 1, \ldots, n$ if $n \le r$, $y = 0, 1, \ldots, r$ if $n > r$ | $\frac{nr}{N}$ | $n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$ | |
| Poisson | $p(y) = \frac{\lambda^y e^{-\lambda}}{y!};$ $y = 0, 1, 2, \ldots$ | $\lambda$ | $\lambda$ | $\exp[\lambda(e^t - 1)]$ |
| Negative binomial | $p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r};$ $y = r, r+1, \ldots$ | $\frac{r}{p}$ | $\frac{r(1-p)}{p^2}$ | $\left[\frac{pe^t}{1-(1-p)e^t}\right]^r$ |