7

Moment Generating Functions and Sums of Independent Random Variables

7.1 Moment Generating Functions

The purpose of this chapter is to introduce moment generating functions (mgf). We have two applications in mind that will be covered in the next section. We will compute the distribution of some sums of independent random variables and we will indicate how moment generating functions may be used to prove the Central Limit Theorem. We start with their definition.

Moment Generating Functions

The moment generating function of a random variable X is defined as

$$M_X(t) = E(e^{tX}).$$

In particular, if X is a discrete random variable, then

$$M_X(t) = \sum_k e^{tk} P(X = k).$$

If X is a continuous random variable and has a density f, then

$$M_X(t) = \int e^{tx} f(x)dx.$$

Note that an mgf is not necessarily defined for all t (because of convergence problems of the series or of the generalized integral). But it is useful even if it is defined only on a small interval. We start by computing some examples.

R. B. Schinazi, *Probability with Statistical Applications*

Example 1. Consider a binomial random variable S with parameters n and p. Compute its mgf.

We have that

$$\begin{aligned} M_S(t) &= E(e^{tS}) = \sum_{k=0}^{n} e^{tk} P(S = k) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (e^t p)^k (1-p)^{n-k}. \end{aligned}$$

We now use the binomial Theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

with $x = e^t p$ and $y = (1-p)$ to get

$$M_S(t) = (pe^t + 1 - p)^n \text{ for all } t.$$

Example 2. Let N be a Poisson random variable with mean λ. We have

$$M_N(t) = E(e^{tN}) = \sum_{k=0}^{\infty} e^{tk} P(N = k) = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} e^{-\lambda} \frac{(e^t \lambda)^k}{k!}.$$

Recall that

$$e^x = \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

We use this power series expansion with $x = e^t \lambda$ to get

$$M_N(t) = e^{-\lambda} \exp(e^t \lambda) = \exp(\lambda(-1 + e^t)) \text{ for all } t.$$

We now give an example of computation of an mgf for a continuous random variable.

Example 3. Assume X is exponentially distributed with rate λ. Its mgf is

$$M_X(t) = E(e^{tX}) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_0^{\infty} \lambda e^{(t-\lambda)x} dx.$$

Note that the preceding improper integral is convergent only if $t - \lambda < 0$. In that case we get

$$M_X(t) = \frac{\lambda}{\lambda - t} \text{ for } t < \lambda.$$

The moment generating functions get their name from the following property.

Moments of a Random Variable

Let X be a random variable. The expectation $E(X^k)$ is called the kth moment of X. If X has a moment generating function M_X defined on some interval $(-r, r)$ for $r > 0$, then all the moments of X exist and

$$E(X^k) = M_X^{(k)}(0)$$

where $M_X^{(k)}$ designates the kth derivative of M_X.

Example 4. We will use the formula above to compute the moments of the Poisson distribution. Let N be a Poisson random variable with mean λ. Then M_N is defined everywhere and

$$M_N(t) = \exp(\lambda(-1 + e^t)).$$

Note that the first derivative is

$$M_N'(t) = \lambda e^t \exp(\lambda(-1 + e^t)).$$

Letting $t = 0$ in the formula above yields

$$E(X) = M_N'(0) = \lambda.$$

We now compute the second derivative

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

Thus,

$$E(X^2) = M_N''(0) = \lambda + \lambda^2.$$

Note that

$$Var(X) = E(X^2) - E(X)^2 = \lambda.$$

Example 5. We now compute the mgf of a standard normal distribution. Let Z be a standard normal distribution. We have

$$M_Z(t) = E(e^{Zt}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{zt} e^{-z^2/2} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{zt - z^2/2} dz.$$

Note that we may 'complete the square' to get

$$zt - z^2/2 = -(z - t)^2/2 + t^2/2.$$

Thus,

$$M_Z(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(z-t)^2/2 + t^2/2} dz = e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(z-t)^2/2} dz.$$

Note that $g(z) = \frac{1}{\sqrt{2\pi}} e^{-(z-t)^2/2}$ is the density of a normal distribution with mean t and standard deviation 1. Thus,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(z-t)^2/2} dz = 1$$

and

$$M_Z(t) = e^{t^2/2}.$$

Example 6. We may use Example 5 to compute the moments of a standard normal distribution.

$$M'_Z(t) = te^{t^2/2}.$$

Letting $t = 0$ above we get

$$E(Z) = 0.$$

We have

$$M''_Z(t) = e^{t^2/2} + t^2 e^{t^2/2}.$$

So

$$E(Z^2) = M''_Z(0) = 1.$$

We also compute the third moment

$$M_Z^{(3)}(t) = te^{t^2/2} + 2te^{t^2/2} + t^3 e^{t^2/2}.$$

We get

$$E(Z^3) = M_Z^{(3)}(0) = 0.$$

Example 7. We now use the computation in Example 5 to compute the mgf of a normal random variable X with mean μ and standard deviation σ. First note that the random variable Z defined as

$$Z = \frac{X - \mu}{\sigma}$$

is a standard normal distribution. We have that

$$M_X(t) = M_{\sigma Z + \mu}(t) = E(e^{t(\sigma Z + \mu)}).$$

Observe that $e^{t\mu}$ is a constant with respect to the expectation so

$$M_X(t) = e^{t\mu} E(e^{t\sigma Z}) = M_Z(t\sigma).$$

We now use that $M_Z(t) = e^{t^2/2}$ to get

$$M_X(t) = \exp(t\mu) \exp(t^2\sigma^2/2) = \exp(t\mu + t^2\sigma^2/2).$$

Our next example deals with the Gamma distribution.

Example 8. A random variable X is said to have a Gamma distribution with parameters $r > 0$ and $\lambda > 0$ if its density is

$$f(x) = e^{-\lambda x}\frac{\lambda^r x^{r-1}}{\Gamma(r)} \text{ for } x > 0$$

where

$$\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx.$$

The improper integral above is convergent for all $r > 0$. Moreover, an easy induction proof shows that

$$\Gamma(n) = (n-1)! \text{ for all integers } n \geq 1.$$

Observe that a Gamma random variable with parameters $r = 1$ and λ is an exponential random variable with parameter λ. We now compute its mgf,

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx}e^{-\lambda x}\frac{\lambda^r x^{r-1}}{\Gamma(r)}dx.$$

The preceding improper integral converges only for $t < \lambda$. We divide and multiply the integrand by $(\lambda - t)^r$ to generate another Gamma distribution.

$$M_X(t) = \frac{\lambda^r}{(\lambda - t)^r}\int_0^\infty e^{-(\lambda - t)x}(\lambda - t)^r\frac{x^{r-1}}{\Gamma(r)}dx.$$

But $g(x) = e^{-(\lambda-t)x}(\lambda - t)^r\frac{x^{r-1}}{\Gamma(r)}$ is the density of a Gamma random variable with parameters r and $\lambda - t$. Thus,

$$\int_0^\infty e^{-(\lambda - t)x}(\lambda - t)^r\frac{x^{r-1}}{\Gamma(r)}dx = 1$$

and

$$M_X(t) = \frac{\lambda^r}{(\lambda - t)^r} \text{ for } t < \lambda.$$

Exercises

1. Compute the moment generating function of a geometric random variable with parameter p.

2. Compute the mgf of a uniform random variable on [0,1].

3. Compute the first three moments of a binomial random variable by taking derivatives of its mgf.

4. Compute the first two moments of a geometric random variable by using Exercise 1.

5. Compute the first two moments of a uniform random variable on [0,1] by using Exercise 2.

6. Use the mgf in Example 8 to compute the mean and standard deviation of a Gamma distribution with parameters n and λ.

7. What is the mgf of a normal distribution with mean 1 and standard deviation 2?

8. Use the mgf in Example 7 to compute the first two moments of a normal distribution with mean μ and standard deviation σ.

9. (a) Make a change of variables to show that

$$\int_0^\infty e^{-\lambda x}\lambda^r x^{r-1}dx = \Gamma(r).$$

(b) Show that for all $r > 0$ and $\lambda > 0$,

$$\int_0^\infty e^{-\lambda x}\frac{\lambda^r x^{r-1}}{\Gamma(r)}dx = 1$$

10. A random variable with density

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}$$

is said to be a Chi-square random variable with n degrees of freedom (n is an integer). Find the moment generating function of X.

7.2 Sums of Independent Random Variables

We first summarize the mgf we have computed in Section 7.1.

Random Variable	Moment generating function
Binomial (n, p)	$(pe^t + 1 - p)^n$
Poisson (λ)	$\exp(\lambda(-1 + e^t))$
Exponential (λ)	$\frac{\lambda}{\lambda - t}$ for $t < \lambda$
Normal (μ, σ^2)	$\exp(t\mu + t^2\sigma^2/2)$
Gamma (r, λ)	$\frac{\lambda^r}{(\lambda-t)^r}$ for $t < \lambda$

We will use moment generating functions to show the following important property of normal random variables.

Linear Combination of Independent Normal Random Variables

Assume that $X_1, X_2, \ldots, X_n$ are independent normal random variables with mean μ_i and variance σ_i^2. Let $a_1, a_2, \ldots, a_n$ be a sequence of real numbers. Then

$$a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

is also a normal variable with mean

$$a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

and variance

$$a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2.$$

The remarkable fact here is that a linear combination of independent normal random variables is normal.

Example 1. Assume that in a given population, heights are normally distributed. The mean height for men is 172 cm with SD 5 cm and for women the mean is 165 cm with SD 3 cm. What is the probability that a woman taken at random is taller than a man taken at random?

Let X be the man's height and let Y be the woman's height. We want $P(X < Y) = P(Y - X > 0)$. According to the preceding property $Y - X$ is normally distributed with

$$E(Y - X) = E(Y) - E(X) = 165 - 172 = -7$$

and

$$Var(Y - X) = Var(Y) + Var(X) = 3^2 + 5^2 = 34.$$

We normalize $Y - X$ to get

$$\begin{aligned} P(X < Y) &= P(Y - X > 0) = p\left(\frac{Y - X - (-7)}{\sqrt{34}} > \frac{0 - (-7)}{\sqrt{34}}\right) \\ &= P\left(Z > \frac{7}{\sqrt{34}}\right) = 0.12. \end{aligned}$$

Example 2. Assume that at a certain university, salaries of junior faculty are normally distributed with mean 40,000 and SD 5,000. Assume also that salaries of senior faculty are normally distributed with mean 60,000 and SD 10,000. What is the probability that the salary of a senior faculty member taken at random is at least twice the salary of a junior faculty member taken at random?

Let X be the salary of the junior faculty member and Y be the salary of the senior faculty member. We want $P(Y > 2X)$. We know that $Y - 2X$ is normally distributed. We express all the figures in thousands of dollars to get

$$E(Y-2X) = -20 \text{ and } Var(Y-2X) = Var(Y)+4Var(X) = 10^2+4\times 5^2 = 200.$$

We normalize to get

$$\begin{aligned} P(Y-2X>0) &= P\left(\frac{Y-2X-(-20)}{\sqrt{200}} > \frac{0-(-20)}{\sqrt{200}}\right) \\ &= P\left(Z > \frac{20}{\sqrt{200}}\right) = 0.08. \end{aligned}$$

Before proving that a linear combination of independent normally distributed random variables is normally distributed we need two properties of moment generating functions that we now state.

P1. The moment generating function of a random variable characterizes its distribution. That is, if two random variables X and Y are such that

$$M_X(t) = M_Y(t) \text{ for all } t \text{ in } (-r, r)$$

for some $r > 0$, then X and Y have the same distribution.

P1 is a crucial property. It tells us that if we recognize a moment generating function then we know what the underlying distribution is.

P2. Assume that the random variables $X_1, X_2, \dots, X_n$ are independent and have moment generating functions. Let $S = X_1 + X_2 + \dots + X_n$; then

$$M_S(t) = M_{X_1}(t)M_{X_2}(t)\dots M_{X_n}(t).$$

The proof of P1 involves mathematics that are beyond the scope of this book. For a proof of P2 see P2 in Section 8.3. We now prove that a linear combination of independent normally distributed random variables is normally distributed. Assume that $X_1, X_2, \dots, X_n$ are independent normal random variables with mean μ_i and variance σ_i^2. Let $a_1, a_2, \dots, a_n$ be a sequence of real numbers. We compute the mgf of $a_1X_1 + a_2X_2 + \dots + a_nX_n$. The random variables a_iX_i are independent so by P2 we have

$$M_{a_1X_1+a_2X_2+\dots+a_nX_n}(t) = M_{a_1X_1}(t)M_{a_2X_2}(t)\dots M_{a_nX_n}(t).$$

Note that by definition

$$M_{a_iX_i}(t) = E(e^{ta_iX_i}) = M_{X_i}(a_it).$$

We now use the mgf corresponding to the normal distribution to get

$$M_{a_iX_i}(t) = \exp(a_it\mu_i + a_i^2t^2\sigma_i^2/2).$$

Thus,

$$M_{a_1X_1+a_2X_2+\cdots+a_nX_n}(t) = \exp(a_1t\mu_1 + a_1^2t^2\sigma_1^2/2) \times \cdots \times \exp(a_nt\mu_n + a_n^2t^2\sigma_2^2/2).$$

Therefore,

$$M_{a_1X_1+a_2X_2+\cdots+a_nX_n}(t) = \exp((a_1\mu_1 + \cdots + a_n\mu_n)t + (a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2)t^2/2).$$

This is exactly the mgf of a normal random variable with mean

$$a_1\mu_1 + \cdots + a_n\mu_n$$

and variance

$$a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2.$$

So according to property P1 this shows that $a_1X_1 + \cdots + a_nX_n$ follows a normal distribution with mean and variance given above.

Example 3. Let $T_1, \ldots, T_n$ be i.i.d. exponentially distributed random variables with rate λ. What is the distribution of $T_1 + T_2 + \cdots + T_n$?

We compute the mgf of the sum by using Property P2.

$$M_{T_1+T_2+\cdots+T_n}(t) = M_{T_1}(t)M_{T_2}(t)\ldots M_{T_n}(t) = M_{T_1}^n(t)$$

since all the T_i have all the same distribution and therefore the same mgf. We use the mgf for an exponential random variable with rate λ to get

$$M_{T_1+T_2+\cdots+T_n}(t) = (\frac{\lambda}{\lambda - t})^n.$$

This is not the mgf of an exponential distribution. However, this is the mgf of a Gamma distribution with parameters n and λ. That is, we have the following.

Sum of i.i.d. Exponential Random Variables

Let $T_1, \ldots, T_n$ be i.i.d. exponentially distributed random variables with rate λ. Then $T_1 + T_2 + \cdots + T_n$ has a Gamma distribution with parameters n and λ.

Example 4. Assume that you have two batteries that have an exponential lifetime with mean two hours. As soon as the first battery fails you replace it with a second battery. What is the probability that the batteries will last at least four hours?

The total time, T, the batteries will last is a sum of two exponential i.i.d. random variables. Therefore, T follows a Gamma distribution with parameters $n = 2$ and $\lambda = 1/2$. We use the density of a Gamma distribution (see Example 8 in 7.1 and note that $\Gamma(2) = 1$) to get

$$P(T > 4) = \int_4^\infty \lambda^2 te^{-\lambda t}dt = 3e^{-2} = 0.41$$

where we use an integration by parts to get the second equality.

Example 5. Let X and Y be two independent Poisson random variables with means λ and μ, respectively. What is the distribution of $X+Y$?

We compute the mgf of $X+Y$. By property P2 we have that

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

Thus,

$$M_{X+Y}(t) = \exp(\lambda(-1+e^t)) \times \exp(\mu(-1+e^t)) = \exp((\lambda+\mu)(-1+e^t)).$$

This is the moment generating function of a Poisson random variable with mean $\lambda+\mu$. Thus, by property P1, $X+Y$ is a Poisson random variable with mean $\lambda+\mu$.

We now state the general result.

Sum of Independent Poisson Random Variables

Let $N_1, \ldots, N_n$ be independent Poisson random variables with means $\lambda_1, \ldots, \lambda_n$, respectively. Then,

$$N_1 + N_2 + \cdots + N_n$$

is a Poisson random variable with mean

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Only a few distributions are stable under addition. Normal and Poisson distributions are two of them.

Example 6. Assume that at a given hospital there are on average two births of twins per month and one birth of triplets per year. Assume that both are Poisson random variables. What is the probability that on a given month there are four or more multiple births?

Let N_1 and N_2 be the numbers of births of twins and of triplets in a given month, respectively. Then $N = N_1 + N_2$ is a Poisson random variable with mean $\lambda = 2 + 1/12 = 25/12$. We have that

$$\begin{aligned} P(N \geq 4) = &\quad 1 - P(N=0) - P(N=1) - P(N=2) - P(N=3) \\ = &\quad 1 - e^{-\lambda} - \lambda e^{-\lambda} - \lambda^2 e^{-\lambda}/2 - \lambda^3 e^{-\lambda}/3! = 0.16. \end{aligned}$$

As noted before, when we sum two random variables with the same type of distribution we do not, in general, get the same distribution. Next, we will look at such an example.

Example 7. Roll two fair dice. What is the distribution of the sum?

Let X and Y be the faces shown by the two dice. The random variables X and Y are discrete uniform random variables on $\{1, 2 \dots, 6\}$. Let $S = X + Y$. Note that S must be an integer between 2 and 12. We have that

$$P(S = 2) = P(X = 1; Y = 1) = P(X = 1)P(Y = 1) = 1/36$$

where we use the independence of X and Y to get the second equality. Likewise, we have that

$$P(S = 3) = P(X = 1, Y = 2) + P(X = 2, Y = 1) = 2/36.$$

In general, we have the following formula

$$P(S = n) = \sum_{k=1}^{n-1} P(X = k)P(Y = n - k) \text{ for } n = 2, 3 \dots, 12.$$

We get the following distribution for S.

k	2	3	4	5	6	7	8	9	10	11	12
$P(X = k)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Note that S is not a uniform random variable. In this case using the moment generating function does not help. We could compute the mgf of S but it would not correspond to any distribution we know.

We now state the general form of the distribution of the sum of two independent random variables.

Sum of Two Independent Random Variables

Let X and Y be two discrete independent random variables. The distribution of $X + Y$ may be computed by using the formula

$$P(X + Y = n) = \sum_k P(X = k)P(Y = n - k).$$

If X and Y are independent continuous random variables with densities f and g, then $X + Y$ has density h that may be computed by using the formula

$$h(x) = \int f(y)g(x - y)dy = \int g(y)f(x - y)dy.$$

We now apply the preceding formula to uniform random variables.

Example 8. Let U and V be two independent uniform random variables on [0,1]. The density for both of them is $f(x) = 1$ for x in [0,1]. Let $S = U + V$ and let h be the density of S. We have that

$$h(x) = \int f(y)f(x-y)dy.$$

Note that $f(y) > 0$ if and only if y is in [0,1]. Note also that $f(x-y) > 0$ if and only if $x-y$ is in [0,1], that is y is in $[-1+x, x]$. Thus, $f(y)f(y-x) > 0$ if and only if y is simultaneously in [0,1] and in $[-1+x, x]$. So

$$h(x) = \int_0^x dy = x \text{ if } x \text{ is in } [0, 1]$$

and

$$h(x) = \int_{-1+x}^1 dy = 2 - x \text{ if } x \text{ is in } [1, 2].$$

Observe that the sum of two uniform random variables is not uniform; the density has a triangular shape instead.

Example 9. Let X and Y be two independent exponentially distributed random variables with rates 1 and 2, respectively. What is the density of $X + Y$?

The densities of X and Y are $f(x) = e^{-x}$ for $x > 0$ and $g(x) = 2e^{-2x}$ for $x > 0$, respectively. The density h of the sum $X + Y$ is

$$h(x) = \int f(y)g(x-y).$$

In order for $f(y)g(x-y) > 0$ we need $y > 0$ and $x - y > 0$. Thus,

$$h(x) = \int_0^x e^{-y}2e^{-2(x-y)}dy \text{ for } x > 0.$$

We get

$$h(x) = 2(e^{-x} - e^{-2x}) \text{ for } x > 0.$$

Note that this is not the density of an exponential distribution. If the two rates were the same we would have obtained a Gamma distribution for the sum, but with different rates we get another type of distribution.

Proof of the Central Limit Theorem

We now sketch the proof of the Central Limit Theorem in the particular case where the random variables have moment generating functions (in the general case it is only assumed that the random variables have finite second moments). Let $X_1, X_2, \ldots, X_n$ be a sequence of independent identically distributed random variables with mean μ and variance σ^2. Assume that X_i has an mgf M_{X_i} defined on $(-r, r)$ for some $r > 0$. Let

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

We want to show that the distribution of

$$T = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

approaches the distribution of a standard normal distribution. We start by computing the moment generating function of T.

$$M_T(t) = E(e^{tT}) = E\left(\exp\left(t\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)\right) = E\left(\exp\left(t\sqrt{n}\frac{\bar{X} - \mu}{\sigma}\right)\right).$$

Observe now that

$$\frac{\bar{X} - \mu}{\sigma} = \frac{1}{n}\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma}.$$

Let $Y_i = \frac{X_i - \mu}{\sigma}$ and $S = \sum_{i=1}^n Y_i$. We have that

$$M_T(t) = E\left(\exp\left(t\frac{\sqrt{n}}{n}\sum_{i=1}^{n} Y_i\right)\right) = E\left(\exp\left(t\frac{\sqrt{n}}{n}S\right)\right) = M_S\left(\frac{t}{\sqrt{n}}\right).$$

Since the Y_i are independent we get by P2 that

$$M_T(t) = M_Y\left(\frac{t}{\sqrt{n}}\right)^n.$$

We now write a third degree Taylor expansion for M_Y.

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = M_Y(0) + \frac{t}{\sqrt{n}}M'_Y(0) + \frac{t^2}{2n}M''_Y(0) + \frac{t^3}{6n^{3/2}}M'''_Y(s)$$

for some s in $(0, \frac{t}{\sqrt{n}})$. Since the Y_i are standardized we have that

$$M'_Y(0) = E(Y) = 0 \text{ and } M''_Y(0) = E(Y^2) = Var(Y) = 1.$$

We also have (for any random variable) that $M_Y(0) = 1$. Thus,

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + \frac{t^3}{6n^{3/2}}M'''_Y(s).$$

We have that

$$\ln(M_T(t)) = \ln\left(M_Y\left(\frac{t}{\sqrt{n}}\right)^n\right) = n\ln\left(M_Y\left(\frac{t}{\sqrt{n}}\right)\right) = n\ln\left(1+\frac{t^2}{2n}+\frac{t^3}{6n^{3/2}}M'''_Y(s)\right).$$

By writing that the derivative of $\ln(1+x)$ at 0 is 1 we get

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=1.$$

Let $x_n=\frac{t^2}{2n}+\frac{t^3}{6n^{3/2}}M_Y'''(s)$ and note that x_n converges to 0 as n goes to infinity. We have

$$n\ln\left(M_Y\left(\frac{t}{\sqrt{n}}\right)\right)=n\ln(1+x_n)=nx_n\frac{\ln(1+x_n)}{x_n}.$$

Note that $M_Y'''(s)$ converges to $M_Y'''(0)$ as n goes to infinity. Thus,

$$\lim_{n\to\infty}nx_n=t^2/2 \text{ and } \lim_{n\to\infty}\frac{\ln(1+x_n)}{x_n}=1.$$

Therefore,

$$\lim_{n\to\infty}\ln(M_T(t))=t^2/2 \text{ and } \lim_{n\to\infty}M_T(t)=e^{t^2/2}.$$

That is, this computation shows that the mgf of $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ converges to the moment generating function of a standard normal distribution. A rather deep result of probability theory called Levy's Continuity Theorem shows that this is enough to prove that the distribution of $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ converges to the distribution of a standard normal random variable. This concludes the sketch of the proof of the CLT.

Exercises

1. The weight of a manufactured product is normally distributed with mean 5 kg and SD 0.1 kg.

(a) Take two items at random, what is the probability that they have a weight difference of at least 0.3 kg?

(b) Take three items at random. What is the probability that the sum of the three weights is less than 14 kg?

2. Consider X a binomial random variable with parameters $n=10$ and p. Let Y be independent of X and be a binomial random variable with $n=15$ and p. Let $S=X+Y$.

(a) Find the mgf of S.

(b) What is the distribution of S?

3. Let X be normally distributed with mean 10 and SD 1. Let $Y=2X-30$.

(a) Compute the mgf of Y.

(b) Use (a) to show that Y is normally distributed and to find the mean and SD of Y.

4. Let X be the number of students from University A that get into Medical School at University B. Let Y be the number of students from University A that get into

Law School at University B. Assume that X and Y are two independent Poisson random variables with means 2 and 3, respectively. What is the probability that $X + Y$ is larger than 5?

5. Assume that 6-year old weights are normally distributed with mean 20 kg and SD 3 kg. Assume that male adults' weights are normally distributed with mean 70 kg and SD 6 kg. What is the probability that the sum of the weights of three children is larger than an adult's weight?

6. Assume you roll a die three times; you win each time you get a 6. Assume you toss a coin twice; you win each time heads comes up. Compute the distribution of your total number of wins.

7. Find the density of a sum of three independent uniform random variables on $[0,1]$. You may use the result for the sum of two uniform random variables in Example 8.

8. Let X and Y be two independent geometric random variables with the same probability of success p. Find the distribution of $X + Y$.

9. (a) Use moment generating functions to show that if X and Y are independent binomial random variables with parameters n and p, and m and p, respectively, then $X + Y$ is also a binomial random variable.

(b) If the probabilities of success are distinct for X and Y, is $X + Y$ a binomial random variable?