Moment generating function I

For any random variable X and for any t we write:

$m_X(t) = Ee^{tX} = \sum_{x \in R_X} e^{tx} f_X(x)$.

We define a family of functions $g_t(x) = e^{tx}, t \in (a, b), a < 0 < b$, and compute the expectations $Eg_t(X)$.

Moment generating function II

Note that the first derivative *with respect to t* is:

$m'_X(t) = \sum_{x \in R_X} x e^{tx} f_X(x)$, so that $m'_X(0) = EX$.

The second derivative is:

$m''_X(t) = \sum_{x \in R_X} x^2 e^{tx} f_X(x)$, so that $m''_X(0) = EX^2$

The r'th derivative is

$m^{(r)}_X(t) = \sum_{x \in R_X} x^r e^{tx} f_X(x)$, so that $m^{(r)}_X(0) = EX^r$.

$m_X(t)$ is called the *moment generating function* of X.

Moment generating function III

Moment generating functions are useful because of the following fact:
Two random variables with the same moment generating function have the same distribution.

Examples:
Bernoulli random variable:

$$m(t) = (e^{0\cdot t} \cdot (1-p) + e^{1\cdot t} p) = 1 - p + pe^t, t \in (-\infty, \infty)$$
$$EX = m'(0) = p, E(X^2) = m''(0) = p.$$

Binomial distribution

$$
\begin{aligned}
m(t) &= \sum_{n=0}^{N} e^{tn} \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=0} \binom{N}{n} (pe^t)^n (1-p)^{N-n} =
\end{aligned}
$$

Binomial distribution

$$
\begin{aligned}
m(t) &= \sum_{n=0}^{N} e^{tn} \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=0} \binom{N}{n} (pe^t)^n (1-p)^{N-n} = [pe^t + (1-p)]^N
\end{aligned}
$$

Binomial distribution

$$
\begin{aligned}
m(t) &= \sum_{n=0}^{N} e^{tn} \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=0} \binom{N}{n} (pe^t)^n (1-p)^{N-n} = [pe^t + (1-p)]^N
\end{aligned}
$$

$$m'(t) = N[pe^t + (1-p)]^{N-1} pe^t$$
$$m''(t) = N(N-1)[pe^t+(1-p)]^{N-2} p^2 e^{2t} + N[pe^t+(1-p)]^{N-1} pe^t.$$

Binomial distribution

$$
\begin{aligned}
m(t) &= \sum_{n=0}^{N} e^{tn} \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=0} \binom{N}{n} (pe^t)^n (1-p)^{N-n} = [pe^t + (1-p)]^N
\end{aligned}
$$

$$
\begin{aligned}
&m'(t) = N[pe^t + (1-p)]^{N-1} pe^t \\
&m''(t) = N(N-1)[pe^t + (1-p)]^{N-2} p^2 e^{2t} + N[pe^t + (1-p)]^{N-1} pe^t.
\end{aligned}
$$

$EX = m'(0) = Np.$

$E(X^2) = m''(0) = N(N-1)p^2 + Np = Np(1-p) + (Np)^2.$

$Var(X) = E(X^2) - (EX)^2 = Np(1-p) + (Np)^2 - (Np)^2 = Np(1-p).$

Central limit theorem explanation

This is a mathematical fact, which we will explain using moment generating functions.
First two important facts:

- The moment generating function for a Normal random variable Y with mean μ and variance σ^2:
  $m_Y(t) = e^{t\mu + t^2\sigma^2/2}$.
  Note: To be consistent with the book we write $N(\mu, \sigma)$ for the Normal distribution with mean μ and variance σ^2.
- If two random variables have the same moment generating function then they have the same distribution.

Central Limit Theorem, I.I.D variables

Let $X_1, \ldots, X_N$ be independent with the same distribution $f_X(x)$. Think of sampling with replacement from a box with numbers distributed as f_X.
Let $\mu = EX_n$, and $\sigma^2 = Var(X_n)$.

As before we have the sum $S_N = \sum_{n=1}^{N} X_n$.

.

Central Limit Theorem, I.I.D variables

Let $X_1, \ldots, X_N$ be independent with the same distribution $f_X(x)$.
Think of sampling with replacement from a box with numbers distributed as f_X.
Let $\mu = EX_n$, and $\sigma^2 = Var(X_n)$.

As before we have the sum $S_N = \sum_{n=1}^{N} X_n$.

Standardized S_N: centered by $E(S_N)$ and scaled by $SD(S_N)$:

$$Z_N = \frac{S_N - ES_N}{SD(S_N)} \qquad .$$

Central Limit Theorem, I.I.D variables

Let $X_1, \ldots, X_N$ be independent with the same distribution $f_X(x)$.
Think of sampling with replacement from a box with numbers distributed as f_X.
Let $\mu = EX_n$, and $\sigma^2 = Var(X_n)$.

As before we have the sum $S_N = \sum_{n=1}^{N} X_n$.

Standardized S_N: centered by $E(S_N)$ and scaled by $SD(S_N)$:

$$Z_N = \frac{S_N - ES_N}{SD(S_N)} = \frac{S_N - N\mu}{\sqrt{N}\sigma} \qquad .$$

Central Limit Theorem, I.I.D variables

Let $X_1, \ldots, X_N$ be independent with the same distribution $f_X(x)$.
Think of sampling with replacement from a box with numbers distributed as f_X.
Let $\mu = EX_n$, and $\sigma^2 = Var(X_n)$.

As before we have the sum $S_N = \sum_{n=1}^{N} X_n$.

Standardized S_N: centered by $E(S_N)$ and scaled by $SD(S_N)$:

$$Z_N = \frac{S_N - ES_N}{SD(S_N)} = \frac{S_N - N\mu}{\sqrt{N}\sigma} = \frac{\sum_{n=1}^{N}(X_n - \mu)}{\sqrt{N}\sigma} = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Y_n.$$

where $Y_n = (X_n - \mu)/\sigma$, is the standardized version of X_n.

Central Limit Theorem, I.I.D variables

Y_n is standardized version of X_n so $EY_n = 0$ and

$EY_n^2 = Var(Y_n) = Var(X_n)/\sigma^2 = 1.$

Let's do a Taylor expansion of the moment generating function of Y_n around 0. for small t

$$m_{Y_n}(t) \sim m_{Y_n}(0) + tm'_{Y_n}(0) + t^2/2m''_{Y_n}(0).$$

We know that $m_{Y_n}(0) = 1$ and $m'_{Y_n}(0) = EY_n = 0$ and $m''_{Y_n}(0) = EY_n^2 = 1$, so

Central Limit Theorem, I.I.D variables

$(*) \quad m_{Y_n}(t) \sim 1 + \frac{1}{2}t^2$, and recall: $Z_n = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Y_n$.

Central Limit Theorem, I.I.D variables

$(*)\quad m_{Y_n}(t) \sim 1 + \frac{1}{2}t^2$, and recall: $Z_n = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Y_n$.

So $\qquad m_{Z_N}(t) = Ee^{\frac{t}{\sqrt{N}} \sum_{n=1}^{N} Y_n}$

Central Limit Theorem, I.I.D variables

$(*) \quad m_{Y_n}(t) \sim 1 + \frac{1}{2}t^2$, and recall: $Z_n = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Y_n$.

$$\begin{aligned} \text{So} \qquad m_{Z_N}(t) &= E e^{\frac{t}{\sqrt{N}} \sum_{n=1}^{N} Y_n} \\ \text{Independence} \qquad &= \prod_{n=1}^{N} E e^{\frac{t}{\sqrt{N}} Y_n} = \prod_{n=1}^{N} m_{Y_n}(t/\sqrt{N}) \end{aligned}$$

Central Limit Theorem, I.I.D variables

$$(*) \quad m_{Y_n}(t) \sim 1 + \tfrac{1}{2}t^2, \text{ and recall: } Z_n = \tfrac{1}{\sqrt{N}} \textstyle\sum_{n=1}^{N} Y_n.$$

$$\begin{aligned}
\text{So} \qquad m_{Z_N}(t) &= E e^{\frac{t}{\sqrt{N}} \sum_{n=1}^{N} Y_n} \\
\text{Independence} \qquad &= \prod_{n=1}^{N} E e^{\frac{t}{\sqrt{N}} Y_n} = \prod_{n=1}^{N} m_{Y_n}(t/\sqrt{N}) \\
\text{By } * \qquad &\sim \left[1 + \frac{1}{2} \left(\frac{t}{\sqrt{N}} \right)^2 \right]^N
\end{aligned}$$

Central Limit Theorem, I.I.D variables

$(*) \quad m_{Y_n}(t) \sim 1 + \frac{1}{2}t^2$, and recall: $Z_n = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Y_n$.

$$
\begin{aligned}
\text{So} \qquad m_{Z_N}(t) &= Ee^{\frac{t}{\sqrt{N}} \sum_{n=1}^{N} Y_n} \\
\text{Independence} \qquad &= \prod_{n=1}^{N} Ee^{\frac{t}{\sqrt{N}} Y_n} = \prod_{n=1}^{N} m_{Y_n}(t/\sqrt{N}) \\
\text{By } * \qquad &\sim \left[1 + \frac{1}{2} \left(\frac{t}{\sqrt{N}} \right)^2 \right]^N \\
&= \left[1 + \frac{1}{2} \frac{t^2}{N} \right]^N \underset{N \to \infty}{\longrightarrow} e^{\frac{1}{2}t^2}.
\end{aligned}
$$

Central Limit Theorem, I.I.D variables

$$(*) \quad m_{Y_n}(t) \sim 1 + \tfrac{1}{2}t^2, \text{ and recall: } Z_n = \tfrac{1}{\sqrt{N}} \textstyle\sum_{n=1}^N Y_n.$$

$$\begin{aligned}
\text{So} \qquad m_{Z_N}(t) &= E e^{\frac{t}{\sqrt{N}} \sum_{n=1}^N Y_n} \\
\text{Independence} \quad &= \prod_{n=1}^N E e^{\frac{t}{\sqrt{N}} Y_n} = \prod_{n=1}^N m_{Y_n}(t/\sqrt{N}) \\
\text{By } * \quad &\sim \left[1 + \frac{1}{2}\left(\frac{t}{\sqrt{N}}\right)^2\right]^N \\
&= \left[1 + \frac{1}{2}\frac{t^2}{N}\right]^N \underset{N\to\infty}{\longrightarrow} e^{\frac{1}{2}t^2}.
\end{aligned}$$

Which is the moment generating function of the standard normal distribution, i.e. with mean 0 and variance 1.

Central Limit Theorem, I.I.D variables, Version 1

Conclusion, version 1: As N gets larger the distribution of Z_N - the standardized sum of N i.i.d random variables with mean μ and variance σ^2 gets closer and closer to $N(0, 1)$.

Central Limit Theorem, I.I.D variables, Version 2

Version 2: The distribution of $\overline{X}$ is approximately $N(\mu, \frac{\sigma}{\sqrt{N}})$.

Central Limit Theorem, I.I.D Normal variables

Special case: X_n are draws from a population that has a normal distribution (in the first place).

If $f_{X_n} = N(\mu, \sigma)$, then $f_{\overline{X}} = N(\mu, \frac{\sigma}{\sqrt{N}})$.

This is exact, it is not an approximation.

Why? Use moment generating functions:

The MGF of X_n is $m_{X_n}(t) = e^{t\mu + t^2\sigma^2/2}$

Central Limit Theorem, I.I.D Normal variables

$$m_{\overline{X}}(t) = Ee^{\frac{t}{N}\sum_{n=1}^{N} X_n} = E\prod_{n=1}^{N} e^{\frac{t}{N}X_n}$$

Central Limit Theorem, I.I.D Normal variables

$$m_{\overline{X}}(t) = Ee^{\frac{t}{N}\sum_{n=1}^{N} X_n} = E\prod_{n=1}^{N} e^{\frac{t}{N}X_n}$$

$$\text{Independence: } = \prod_{n=1}^{N} Ee^{\frac{t}{N}X_n}$$

Central Limit Theorem, I.I.D Normal variables

$$
\begin{aligned}
m_{\overline{X}}(t) = & E e^{\frac{t}{N}\sum_{n=1}^{N} X_n} = E\prod_{n=1}^{N} e^{\frac{t}{N}X_n} \\
\text{Independence: } = & \prod_{n=1}^{N} E e^{\frac{t}{N}X_n} \\
\text{MGF of Normal at } t/N: \quad = & \left(e^{\frac{t^2}{N^2}\frac{\sigma^2}{2}+\frac{t}{N}\mu}\right)^N = e^{t^2\frac{\sigma^2}{2N}+t\mu},
\end{aligned}
$$

Central Limit Theorem, I.I.D Normal variables

$$
\begin{aligned}
m_{\overline{X}}(t) =& E e^{\frac{t}{N}\sum_{n=1}^{N} X_n} = E\prod_{n=1}^{N} e^{\frac{t}{N}X_n} \\
\text{Independence:}\ =& \prod_{n=1}^{N} E e^{\frac{t}{N}X_n} \\
\text{MGF of Normal at } t/N: \quad =& \left(e^{\frac{t^2}{N^2}\frac{\sigma^2}{2}+\frac{t}{N}\mu}\right)^N = e^{t^2\frac{\sigma^2}{2N}+t\mu},
\end{aligned}
$$

which is the MGF of $N(\mu, \frac{\sigma}{\sqrt{N}})$.

Conclusion: The average of N independent normal $N(\mu, \sigma)$ random variables is also normal with mean μ and variance σ^2/N.