Probability and Stochastic Processes I - Lecture 21

Michael Evans
University of Toronto
http://www.utstat.utoronto.ca/mikevans/stac62/STAC622023.html

2023

III.9 Generating Functions and the Characteristic Function

- consider a sequence $\{a_n : n \in \mathbb{N}_0\}$ of real numbers, then the *generating function* of the sequence is defined by

$$G(t) = \sum_{i=0}^{\infty} a_i t^i$$

provided the series converges for all $t \in (-h_G, h_G)$ with $h_G > 0$ as then

$$\left.\frac{d^k G(t)}{dt^k}\right|_{t=0} = a_k k!$$

- not all sequences have generating functions (e.g. $a_n = n!$)
- if $G(t) = \sum_{i=0}^{\infty} a_i t^i, H(t) = \sum_{i=0}^{\infty} b_i t^i$ are generating functions, then

$$K(t) = G(t)H(t) = \sum_{i=0}^{\infty} c_i t^i \text{ where } c_i = a_0 b_i + a_1 b_{i-1} + \cdots + a_i b_0$$

is the generating function of $\{c_n : n \in \mathbb{N}_0\}$ where $h_K = \min\{h_G, h_H\}$

Abel's Theorem If $G(t) = \sum_{i=0}^{\infty} a_i t^i$ is finite in $(-1, 1)$ and $\sum_{i=0}^{\infty} a_i$ converges (limit could be ∞), then $\lim_{t\uparrow 1} G(t) = \sum_{i=0}^{\infty} a_i$.
Proof: See a book on Analysis.

Probability Generating Functions

Definition III.9.1 If X is a r.v. such that $P_X(\mathbb{N}_0) = 1$, then $G_X(t) = E(t^X) = \sum_{i=0}^{\infty} P(X = i)t^i$ for $|t| \le 1$ is the *probability generating function* of X. ■

Proposition III.9.1 If $G_X(t) = G_Y(t)$ for all $t \in (-h, h)$ for some $h > 0$, then X and Y have the same probability distribution.

Proof: Since $G_X(t) = \sum_{i=0}^{\infty} P(X = i)t^i$ for $|t| \le 1$, then for $|t| < 1$

$$\frac{1}{k!}\frac{d^k G_X(t)}{dt^k}\bigg|_{t=0} = P(X = k) = \frac{1}{k!}\frac{d^k G_Y(t)}{dt^k}\bigg|_{t=0} = P(Y = k). \blacksquare$$

- so G_X completely specifies the distribution of X

Proposition III.9.2 (i) If X, Y are stat. ind. r.v.'s with pgf's G_X, G_Y, then $G_{X+Y}(t) = G_X(t)G_Y(t)$.

(ii) If X has pgf G_X and the *k-th factorial moment*

$$\mu_{[k]} = E(X(X-1)\cdots(X-k+1)) = \sum_{i=k}^{\infty} i(i-1)\cdots(i-k+1)P(X=i)$$

of X exists then $\lim_{t\uparrow 1} \frac{d^k G_X(t)}{dt^k} = \mu_{[k]}$.

(iii) (*Compound distributions*) If the r.v.'s $\{X_i : i = 1, 2, \ldots\}$ are i.i.d. with pgf G_X, stat. ind. of N with pgf G_N, then $Y = \sum_{i=1}^{N} X_i$ has pgf $G_Y(t) = G_N(G_X(t))$.

Proof: (i)

$$G_{X+Y}(t) = E(t^{X+Y}) = E(t^X t^Y) \overset{\text{ind}}{=} E(t^X)E(t^Y) = G_X(t)G_Y(t).$$

(ii) When $|t| < 1$, then

$$\frac{d^k G_X(t)}{dt^k} = \frac{d^k}{dt^k} \sum_{i=0}^{\infty} P(X = i) t^i = \sum_{i=k}^{\infty} i(i-1) \cdots (i-k+1) P(X = i) t^{i-k}$$

is finite and by Abel's Thm

$$\lim_{t \uparrow 1} \sum_{i=k}^{\infty} i(i-1) \cdots (i-k+1) P(X = i) t^{i-k}$$
$$= \sum_{i=k}^{\infty} i(i-1) \cdots (i-k+1) P(X = i) = \mu_{[k]}.$$

(iii)

$$G_Y(t) = E(t^Y) = E\left(t^{\sum_{i=1}^{N} X_i}\right) = E\left(\prod_{i=1}^{N} t^{X_i}\right)$$
$$\overset{TTE}{=} E\left(E\left(\prod_{i=1}^{N} t^{X_i} \,\middle|\, N\right)\right) = \sum_{n=1}^{\infty} P(N = n) E\left(\prod_{i=1}^{n} t^{X_i}\right)$$
$$\overset{(i)}{=} \sum_{n=1}^{\infty} P(N = n) G_X^n(t) = G_N(G_X(t)). \blacksquare$$

Example III.9.1 *Poisson*

- if $X \sim \text{Poisson}(\lambda)$ with $\lambda > 0$, then

$$p_X(i) = \frac{\lambda^i}{i!}e^{-\lambda} \text{ for } i = 0, 1, 2, \ldots$$

and

$$G_X(t) = E(t^X) = \sum_{i=0}^{\infty} t^i \frac{\lambda^i}{i!}e^{-\lambda} = e^{-\lambda}\sum_{i=0}^{\infty}\frac{(t\lambda)^i}{i!} = e^{-\lambda}e^{t\lambda} = e^{\lambda(t-1)}$$

- so if $X \sim \text{Poisson}(\lambda_1)$ ind. of $Y \sim \text{Poisson}(\lambda_2)$, then

$$G_{X+Y}(t) = G_X(t)G_Y(t) = e^{\lambda_1(t-1)}e^{\lambda_2(t-1)} = e^{(\lambda_1+\lambda_2)(t-1)}$$

and therefore $X + Y$ Poisson$(\lambda_1 + \lambda_2)$

- if $X \sim \text{Poisson}(\lambda)$, then since $\sum_{i=0}^{\infty}\frac{(t\lambda)^i}{i!}$ converges for all $t \in R^1$, then $\mu_{[k]}$ is finite for all k and

$$
\begin{aligned}
\mu_1 &= \mu_{[1]} = \lim_{t \uparrow 1} \frac{dG_X(t)}{dt} = \lim_{t \uparrow 1} \lambda e^{\lambda(t-1)} = \lambda \\
\mu_{[2]} &= \lim_{t \uparrow 1} \frac{d^2 G_X(t)}{dt^2} = \lim_{t \uparrow 1} \lambda^2 e^{\lambda(t-1)} = \lambda^2 \\
Var(X) &= \mu_{[2]} - \mu_{[1]}(\mu_{[1]} - 1) = \lambda^2 - \lambda(\lambda - 1) = \lambda \ \blacksquare
\end{aligned}
$$

Exercise III.9.1 If $X \sim$ Bernoulli(p), then find $G_X(t)$ and use this to obtain the pgf for a binomial(n, p) distribution.

Exercise III.9.2 If $X \sim$ Geometric(p), then find $G_X(t)$ and use this to obtain the mean and variance of X.

Exercise III.9.3 If $N \sim$ Poisson(λ) independent of $X_1, X_2, \ldots \sim -1 + 2$Bernoulli$(p)$ and $Y = \sum_{i=1}^{N} X_i$, determine $E(Y)$.

Moment Generating Function and Characteristic Function

Definition III.9.2 (i) If $\mathbf{X} \in R^k$ is a random vector, then $m_{\mathbf{X}}(\mathbf{t}) = E(\exp(\mathbf{t}'\mathbf{X}))$ is the *moment generating function* of $\mathbf{X}$ provided the expectation is finite for all $\mathbf{t} \in B_h(\mathbf{0})$, for some $h > 0$. (ii) The characteristic function of $\mathbf{X}$ is given by $c_{\mathbf{X}}(\mathbf{t}) = E(\exp(i\mathbf{t}'\mathbf{X}))$ for all $\mathbf{t} \in R^k$. ■

- $m_{\mathbf{X}}$ may not exist but since $e^{ix} = \cos x + i \sin x$ and $|\cos x| \leq 1, |\sin x| \leq 1$ and

$$\begin{aligned} & E(|\exp(i\mathbf{t}'\mathbf{X})|) \\ = \ & E(|\cos(\mathbf{t}'\mathbf{X}) + i \sin(\mathbf{t}'\mathbf{X})|) \leq E(|\cos(\mathbf{t}'\mathbf{X})|) + E(|\sin(\mathbf{t}'\mathbf{X})|) \leq 2 \end{aligned}$$

so $c_{\mathbf{X}}(\mathbf{t}) = E(\cos(\mathbf{t}'\mathbf{X})) + iE(\sin(\mathbf{t}'\mathbf{X}))$ always exists (may be complex-valued)

- if $P_{\mathbf{X}}(B) = P_{\mathbf{X}}(-B)$ then $P_{\mathbf{X}}(\mathbf{t}'\mathbf{X} \leq x) = P_{\mathbf{X}}(\mathbf{t}'\mathbf{X} \geq -x)$ and $\mathbf{t}'\mathbf{X}$ has a probability distribution symmetric about 0 and since $\sin(-x) = -\sin(x)$, this implies $E(\sin(\mathbf{t}'\mathbf{X})) = 0$ and $c_{\mathbf{X}}$ is real-valued

Proposition III.9.3 *(Uniqueness)* (i) If $m_{\mathbf{X}}, m_{\mathbf{Y}}$ exist and $m_{\mathbf{X}}(\mathbf{t}) = m_{\mathbf{Y}}(\mathbf{t})$ for all $\mathbf{t} \in B_h(\mathbf{0})$, for some $h > 0$, then $P_{\mathbf{X}} = P_{\mathbf{Y}}$. (ii) If $c_{\mathbf{X}}(\mathbf{t}) = c_{\mathbf{Y}}(\mathbf{t})$ for all $\mathbf{t} \in R^k$ then $P_{\mathbf{X}} = P_{\mathbf{Y}}$.
Proof: Accept.

- so if we know $m_{\mathbf{X}}$ or $c_{\mathbf{X}}$ and we recognize it then we know the distribution of $\mathbf{X}$

- there are inversion results that give expressions for the cdf of $\mathbf{X}$ computed from $m_{\mathbf{X}}$ or $c_{\mathbf{X}}$

Definition III.9.3 If $i_1, \ldots, i_k \in \mathbb{N}_0$, then $(i_1, \ldots, i_k)$-th mixed moment of random vector $\mathbf{X} \in R^k$ is defined by

$$\mu_{i_1,\ldots,i_k} = E(X_1^{i_1} \cdots X_k^{i_k})$$

whenever this expectation exists. ■

Proposition III.9.4 If $i_1 \le j_1, \ldots, i_k \le j_k$ and $E(|X_1^{j_1} \cdots X_k^{j_k}|) < \infty$ for all $(j_1, \ldots, j_k)$ satisfying $j_1 + \cdots + j_k = j$ then $\mu_{i_1,\ldots,i_k}$ is finite.
Proof: **Exercise III.9.4** Do the case when $k = 2$.

Proposition III.9.5 If $m_{\mathbf{X}}$ exists, then all the moments of $\mathbf{X}$ are finite and

$$\mu_{i_1,\ldots,i_k} = \left.\frac{\partial^k m_{\mathbf{X}}(\mathbf{t})}{\partial^{i_1} t_1 \cdots \partial^{i_k} t_k}\right|_{\mathbf{t}=\mathbf{0}}.$$

Proof: Consider the case when $k = 1$. Then for $t \in B_h(0)$

$$\begin{aligned} m_X(t) &= E(\exp(tX)) = E(I_{\{X\geq 0\}}\exp(tX_+)) + E(I_{\{X<0\}}\exp(-tX_-)) \\ &= m_{X_+}(t) - P(X<0) + m_{X_-}(-t) - P(X\geq 0) < \infty \end{aligned}$$

(since, for example, $P(X_+ = 0) = P(X = 0) + P(X < 0)$) so m_{X_+} and m_{X_-} exist which implies $m_{|X|}(t) = E(\exp(tX_+ + tX_-)) =$ $m_{X_+}(t) - P(X<0) + m_{X_-}(t) - P(X \geq 0) < \infty$ and so $m_{|X|}$ exists. Let

$$\begin{aligned} Y_n &= \sum_{j=0}^{n} \frac{t^j X^j}{j!} \to \sum_{j=0}^{\infty} \frac{t^j X^j}{j!} = \exp(tX) \text{ so} \\ |Y_n| &\leq \sum_{j=0}^{n} \frac{|t|^j |X|^j}{j!} \uparrow \sum_{k=0}^{\infty} \frac{|t|^j |X|^j}{j!} = \exp(|t||X|). \end{aligned}$$

Since $m_{|X|}$ exists $E(|X|^k) \leq \frac{k!}{|t|^k} m_{|X|}(|t|) < \infty$ and so all moments of X are finite. Furthermore, by DCT

$$\lim_{n\to\infty} E(Y_n) \to \sum_{j=0}^{\infty} \frac{t^j \mu_j}{j!} = m_X(t)$$

which implies

$$\mu_j = \left. \frac{d^j m_X(t)}{dt^j} \right|_{t=0}.$$

For the general case put $\mathbf{Z} = (|X_1|, \ldots, |X_k|)$ and a similar argument shows that $m_{\mathbf{Z}}$ exists. Put

$$\begin{aligned} Y_n &= \sum_{j=0}^{n} \frac{(t_1 X_1 + \cdots + t_k X_k)^j}{j!} \\ &= \sum_{j=0}^{n} \frac{1}{j!} \sum_{\substack{i_1 \geq 0 \ldots i_k \geq 0 \\ i_1 + \cdots + i_k = j}} \binom{j}{i_1 \ldots i_k} t_1^{i_1} \cdots t_k^{i_k} X_1^{i_1} \cdots X_k^{i_k} \\ |Y_n| &\leq \exp(|t_1||X_1| + \cdots + t_k |X_k|) \end{aligned}$$

which implies $\mu_{i_1,\ldots,i_k}$ is finite and by DCT

$$E(Y_n) \to \sum_{j=0}^{\infty} \sum_{\substack{i_1 \geq 0 \ldots i_k \geq 0 \\ i_1 + \cdots + i_k = j}} \frac{t_1^{i_1} \cdots t_k^{i_k}}{i_1! \cdots i_k!} \mu_{i_1,\ldots,i_k} = m_{\mathbf{X}}(\mathbf{t}). \blacksquare$$

Proposition III.9.6 If $m_{\mathbf{X}}$ exists, then $c_{\mathbf{X}}(\mathbf{t}) = m_{\mathbf{X}}(i\mathbf{t})$.
Proof: Accept.

Proposition III.9.7 If $\mathbf{X}, \mathbf{Y} \in R^k$ are stat. ind. with mgf's $m_{\mathbf{X}}, m_{\mathbf{Y}}$ (cf′s $c_{\mathbf{X}}, c_{\mathbf{Y}}$) then $\mathbf{X} + \mathbf{Y}$ has mgf $m_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) = m_{\mathbf{X}}(\mathbf{t})m_{\mathbf{Y}}(\mathbf{t})$ when $m_{\mathbf{X}}(\mathbf{t})$ and $m_{\mathbf{Y}}(\mathbf{t})$ are finite and cf $c_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) = c_{\mathbf{X}}(\mathbf{t})c_{\mathbf{Y}}(\mathbf{t})$.
Proof:

$$\begin{aligned} c_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) &= E(\exp(i\mathbf{t}'(\mathbf{X}+\mathbf{Y}))) = E(\exp(i\mathbf{t}'\mathbf{X})\exp(i\mathbf{t}'\mathbf{Y})) \\ &= E(\exp(i\mathbf{t}'\mathbf{X}))\mathbf{E}(\exp(i\mathbf{t}'\mathbf{Y})) = c_{\mathbf{X}}(\mathbf{t})c_{\mathbf{Y}}(\mathbf{t}). \blacksquare \end{aligned}$$

Example III.9.2 *Normal*

- suppose $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \Sigma)$, then $\mathbf{X} = \boldsymbol{\mu} + \Sigma^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N_k(\mathbf{0}, I)$ so $Z_1, \ldots, Z_k \overset{i.i.d.}{\sim} N(0,1)$ and

$$\begin{aligned} m_{\mathbf{Z}}(\mathbf{t}) &= E(\exp(\mathbf{t}'\mathbf{Z})) = E(\exp(t_1 Z_1 + \cdots + t_k Z_k)) \\ &= E\left(\prod_{i=1}^{k} \exp(t_i Z_i)\right) \overset{i.i.d.}{=} \prod_{i=1}^{k} E(\exp(t_i Z_i)) = \prod_{i=1}^{k} m_Z(t_i) \text{ where} \\ m_Z(t) &= \int_{-\infty}^{\infty} \exp(tz) \frac{1}{\sqrt{2\pi}} \exp(-z^2/2)\, dz \\ &= \exp(t^2/2) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-(z-t)^2/2)\, dz = \exp(t^2/2) \end{aligned}$$

so $m_{\mathbf{Z}}(\mathbf{t}) = \exp(\mathbf{t}'\mathbf{t}/2)$ and

$$\begin{aligned} m_{\mathbf{X}}(\mathbf{t}) &= E(\exp(\mathbf{t}'(\boldsymbol{\mu} + \Sigma^{1/2}\mathbf{Z}))) = \exp(\mathbf{t}'\boldsymbol{\mu}) E(\exp(\mathbf{t}'\Sigma^{1/2}\mathbf{Z})) \\ &= \exp(\mathbf{t}'\boldsymbol{\mu}) E(\exp((\Sigma^{1/2}\mathbf{t})'\mathbf{Z})) = \exp(\mathbf{t}'\boldsymbol{\mu}) \exp(\mathbf{t}'\Sigma\mathbf{t}/2) \\ &= \exp(\mathbf{t}'\boldsymbol{\mu} + \mathbf{t}'\Sigma\mathbf{t}/2) \\ c_{\mathbf{X}}(\mathbf{t}) &= \exp(i\mathbf{t}'\boldsymbol{\mu} - \mathbf{t}'\Sigma\mathbf{t}/2) \text{ using Prop. III.9.6} \end{aligned}$$

- so if $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is a sample from the $N_k(\boldsymbol{\mu}, \Sigma)$ distribution and

$$\mathbf{Y} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i = \text{ sample mean}$$

then

$$\begin{aligned} m_{\mathbf{Y}}(\mathbf{t}) &= E\left(\exp\left(\mathbf{t}'\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\right)\right) = E\left(\prod_{i=1}^{n}\exp\left(\left(\frac{\mathbf{t}}{n}\right)'\mathbf{X}_i\right)\right) \\ &\overset{i.i.d.}{=} \prod_{i=1}^{n} m_{\mathbf{X}}(\mathbf{t}/n) = \exp(\mathbf{t}'\boldsymbol{\mu} + \mathbf{t}'\Sigma\mathbf{t}/2n) \text{ and by Uniqueness} \\ \mathbf{Y} &\sim N_k(\boldsymbol{\mu}, \Sigma/n) \ \blacksquare \end{aligned}$$

Proposition III.9.8 If $\mathbf{X} \in R^k$ is a random vector and $\mathbf{r}'\mathbf{X}$ is normally distributed for every constant $\mathbf{r} \in R^k$, then $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \Sigma)$ for some $(\boldsymbol{\mu}, \Sigma)$.
Proof: We have that $E(\mathbf{r}'\mathbf{X}) = \mathbf{r}'E(\mathbf{X})$ and $Var(\mathbf{r}'\mathbf{X}) = \mathbf{r}'Var(\mathbf{X})\mathbf{r}$ and so put $(\boldsymbol{\mu}, \Sigma) = (E(\mathbf{X}), Var(\mathbf{X}))$. Now

$$m_{\mathbf{r}'\mathbf{X}}(t) = \exp(t\mathbf{r}'\boldsymbol{\mu} + t^2\mathbf{r}'\Sigma\mathbf{r}/2) = m_{\mathbf{X}}(t\mathbf{r})$$

which implies the result. ■

Example III.9.3 *Cauchy*

- suppose $X \sim$ Cauchy, then $E(X)$ does not exist so m_X does not exist
- but using contour integration it can be shown that $c_X(t) = \exp(-|t|)$
- now suppose $X_1, \ldots, X_n$ is a sample from the Cauchy and $Y = \frac{1}{n}\sum_{i=1}^n X_i$
- then

$$c_Y(t) = \prod_{i=1}^n \exp(-|t|/n) = \exp(-|t|)$$

so by Uniqueness $Y \sim$ Cauchy ■

- note that any cf c_X satisfies $c_X(0) = 1$ and by DCT

$$\lim_{t\to 0} c_X(t) = \lim_{t\to 0} E(\cos(tX)) + i \lim_{t\to 0} E(\sin(tX)) = 1$$

so c_X is continuous at 0

- if c_X is also real then $c_X(-t) = E(\cos(-tX)) = E(\cos(tX)) = c_X(t)$ so c_X is symmetric and for any n and $x_1, \ldots, x_n, t_1, \ldots, t_n$

$$\sum_{j=1}^{n} \sum_{k=1}^{n} x_j x_k c_X(t_j - t_k) = E\left(\left|\sum_{j=1}^{n} x_j \exp(it_j X)\right|^2\right) \geq 0$$

- therefore such a c_X can serve as the autocorrelation function of a weakly stationary process

- for any constant a, then $c_X(t) = \exp(-a|t|)$ is such an autocorrelation function as is $c_X(t) = \exp(-a^2|t|)$

Exercise III.9.4 If $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are mut. stat. ind. with $\mathbf{X}_i \sim N_{k_i}(\boldsymbol{\mu}_i, \Sigma_i)$ and $\mathbf{a} \in R^m$, $C_i \in R^{m \times k_i}$ are constant, then determine the distribution of $Y = \mathbf{a} + \sum C_i \mathbf{X}_i$.

Exercise III.9.5 E&R 3.4.13

Exercise III.9.6 E&R 3.4.16

Exercise III.9.7 E&R 3.4.20

Exercise III.9.8 E&R 3.4.29