# STAT 3100 Lecture Note

## Week Three (Sep 20 & 22, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Sep/15 21:02:48

---

# Multivariate normal (MVN) distribution (con'd)

## Marginals of MVN

- Suppose $p$-vector $\mathbf{X} = (X_1, \ldots, X_p)^\top$ and $q$-vector $\mathbf{Y} = (Y_1, \ldots, Y_q)^\top$ are jointly normally distributed. Then, $\mathbf{X}$ and $\mathbf{Y}$ are independent $\Leftrightarrow \operatorname{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}_{p \times q}$.

- If $\mathbf{X}$ is of MVN, then its all margins are still of MVN. The inverse proposition does NOT hold.
    - Cautionary example: Let $Y = XZ$, where $X \sim \mathcal{N}(0, 1)$; $Z$ is independent of $X$ with $\Pr(Z = 1) = \Pr(Z = -1) = .5$. $X$ and $Y$ both turn out to be of standard normal, but they are not jointly normal. (Why?)

---

```r
if (!("plot3D" %in% rownames(installed.packages())))
  install.packages("plot3D")
set.seed(1)
xsize = 1e4L
X = rnorm(xsize)
Z = rbinom(n = xsize, 1, .5)
Y = (2 * Z - 1) * X
# 3d histogram of (X, Y)
plot3D::hist3D(z=table(cut(X, 100), cut(Y, 100)), border = "black")
# plot the support of joint pdf of (X, Y)
plot3D::image2D(z=table(cut(X, 100), cut(Y, 100)), border = "black")
```

# Normal sampling theory (CB Sec. 5.3)

## Stochastic representations for $\chi^2$-, $t$-, and $F$-r.v. (HMC Chp. 3)

- If iid $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$, then
    - $\sum_{i=1}^n X_i^2 \sim \chi^2(n)$ if iid $X_1, \ldots, X_n \sim \mathcal{N}(0, 1)$;
    - $X/\sqrt{Y/n} \sim t(n)$ if $X \sim \mathcal{N}(0, 1)$ and $Y \sim \chi^2(n)$ are independent;
    - $(X/m)/(Y/n) \sim F(m, n)$ if $X \sim \chi^2(m)$ and $Y \sim \chi^2(n)$ are independent.

## Important identities for normal samples

- $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ and $S^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$ are independent
- $n^{1/2}(\bar{X} - \mu)/\sigma \sim \mathcal{N}(0, 1)$

- $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$
- $n^{1/2}(\bar{X} - \mu)/S \sim t(n-1)$

# Taylor series (Def 5.5.20 & Thm 5.5.21 in C&B)

## Taylor series about $x_0 \in \mathbb{R}$ for univariate functions

- Suppose $f$ has derivative of order $n+1$ within an open interval of $x_0$, say $(x_0 - \varepsilon, x_0 + \varepsilon)$ with $\varepsilon > 0$. Then, for $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$,

$$f(x) \approx \sum_{k=0}^{n} \frac{f^{(n)}(x_0)}{k!}(x - x_0)^k = f(x_0) + \sum_{k=1}^{n} \frac{f^{(n)}(x_0)}{k!}(x - x_0)^k,$$

  where $f^{(n)}(x_0) = \frac{\mathrm{d}^n}{\mathrm{d}x^n}f(x)|_{x=x_0}$.
- Called the Maclaurin series if $x_0 = 0$

## Taylor series about $\boldsymbol{x}_0 \in \mathbb{R}^p$ for multivariate functions

Under regularity conditions,

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}_0) + (\boldsymbol{x} - \boldsymbol{x}_0)^\top \nabla f(\boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^\top \mathbf{H}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0),$$

where the gradient $\nabla f(\boldsymbol{x}_0) = [\frac{\partial}{\partial x_1}f(\boldsymbol{x}_0), \cdots, \frac{\partial}{\partial x_p}f(\boldsymbol{x}_0)]^\top \in \mathbb{R}^{p \times 1}$ and the Hessian $\mathbf{H}(\boldsymbol{x}_0) = [\frac{\partial^2}{\partial x_i \partial x_j}f(\boldsymbol{x}_0)]_{p \times p}$.

## Application

- Approximate unknown or complex $f$ with a polynomial
  - $\Delta$-method
  - Asymptotic theory for maximum likelihood estimators
- Moment generating function (mgf): $M_X(t) = \mathrm{E}\{\exp(tX)\} = \sum_{n=0}^{\infty} t^n \mathrm{E}(X^n)/n!$
  - the Maclaurin series of $\exp(tX)$ is $\exp(tX) = \sum_{n=0}^{\infty}(tX)^n/n!$
  - $\Rightarrow \mathrm{E}(X^n) = M_X^{(n)}(0)$

# Generating functions

## Moment generating function (mgf, Sec. 2.3 in C&B)

- Univariate r.v. $X$
  - mgf $M_X(t) = \mathrm{E}\{\exp(tX)\}$ if $\mathrm{E}\exp(tX) < \infty$ for $t$ in a neighborhood of 0; otherwise we say that the mgf does not exist or is undefined.
    * Continuous $X$: $M_X(t) = \int_{-\infty}^{\infty} \exp(tx)f_X(x)\mathrm{d}x$
      · (Two-sided) Laplace transformation of $f_X$
    * Discrete $X$: $M_X(t) = \sum_{\{x:x \in \mathrm{supp}(X)\}} \exp(tx)p_X(x)$
  - $M_{aX+b}(t) = \exp(bt)M_X(at)$

  ---

- Multivariate r.v. $\mathbf{X} = (X_1, \ldots, X_p)^\top \in \mathbb{R}^p$
  - mgf $M_{\mathbf{X}}(\boldsymbol{t})$ is defined as

$$M_{\mathbf{X}}(\boldsymbol{t}) = \mathrm{E}\exp(\boldsymbol{t}^\top \mathbf{X}) = \begin{cases} \int_{\mathbb{R}^p} \exp(\boldsymbol{t}^\top \mathbf{X})f_{\mathbf{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} & \text{continuous } \mathbf{X} \\ \sum_{\{\boldsymbol{x}:\boldsymbol{x} \in \mathrm{supp}(\mathbf{X})\}} \exp(\boldsymbol{t}^\top \mathbf{X})p_{\mathbf{X}}(\boldsymbol{x}) & \text{discrete } \mathbf{X} \end{cases}$$

provided that $\mathrm{E}\exp(\boldsymbol{t}^\top \mathbf{X}) < \infty$ for $\boldsymbol{t} = (t_1, \ldots, t_p)^\top$ in some neighborhood of $\mathbf{0}$; otherwise we say that the mgf does not exist or is undefined.

* $X_1, \ldots, X_p$ are independent $\Rightarrow M_{\mathbf{X}}(\boldsymbol{t}) = \prod_{i=1}^p M_{X_i}(t_i)$

&ndash; $M_{\mathbf{AX}+\boldsymbol{b}}(\boldsymbol{t}) = \exp(\boldsymbol{b}^\top \boldsymbol{t}) M_{\mathbf{X}}(\mathbf{A}^\top \boldsymbol{t})$

---

- Application
  - Characterizing distributions: $M_{\mathbf{X}}(\boldsymbol{t})$ and $M_{\mathbf{Y}}(\boldsymbol{t})$ are both well-defined and equal for all $\boldsymbol{t}$ in a neighborhood of $\mathbf{0} \Leftrightarrow \mathbf{X} \stackrel{d}{=} \mathbf{Y}$
    * Proofs for laws of large numbers and central limit theorems.
  - Computing moments
    * $n$th raw moment $\mu'_n = \mathrm{E}X^n = \sum_{k=0}^n \binom{n}{k} \mu_k (\mu'_1)^{n-k}$
    * $n$th central moment $\mu_n = \mathrm{E}(X - \mathrm{E}X)^n = \sum_{k=0}^n \binom{n}{k}(-1)^{n-k}\mu'_k(\mu'_1)^{n-k}$

## Example Lec6.1

- Find the mgfs of following distributions.
  - $N(\mu, \sigma^2)$.
  - $\mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
  - Cauchy distribution: $f_X(x) = \{\pi(1+x^2)\}^{-1}$, $x \in \mathbb{R}$.

## Other generating functions

- Characteristic function
  - For univariate $X$: $\phi_X(t) = \mathrm{E}\exp(itX)$ for all $t \in \mathbb{R}$
    * Fourier transform of $f_X$
    * Inverse: $f_X(x) = (2\pi)^{-1} \int_{\mathbb{R}} \phi_X(t)\exp(-itx)\mathrm{d}t$
    * $\mu'_n = \mathrm{E}X^n = (-i)^n \phi_X^{(n)}(0)$
  - For Multivariate $\mathbf{X} = (X_1, \ldots, X_p)^\top$: $\phi_{\mathbf{X}}(\boldsymbol{t}) = \mathrm{E}\exp(i\boldsymbol{t}^\top \mathbf{X})$ for all $\boldsymbol{t} \in \mathbb{R}^p$
    * Fourier transform of $f_{\mathbf{X}}$
    * Inverse: $f_{\mathbf{X}}(\boldsymbol{x}) = (2\pi)^{-p} \int_{\mathbb{R}^p} \phi_{\mathbf{X}}(\boldsymbol{t})\exp(-i\boldsymbol{t}^\top \boldsymbol{x})\mathrm{d}\boldsymbol{t}$
  - $\phi_{\mathbf{X}}(\boldsymbol{t}) = \phi_{\mathbf{Y}}(\boldsymbol{t})$ for all $\boldsymbol{t} \in \mathbb{R}^p \Leftrightarrow \mathbf{X} \stackrel{d}{=} \mathbf{Y}$
- Cumulant generating function
  - $K_X(t) = \ln M_X(t) = \sum_{n=0}^\infty \kappa_n t^n / n!$
  - $\kappa_n = K_X^{(n)}(0)$
- Probability-generating function
  - For discrete r.v. $X$ taking values from $\{0, 1, \ldots\}$, $G(z) = \mathrm{E}t^X = \sum_{x=0}^\infty t^x p_X(x)$.
  - $p_X(n) = \Pr(X = n) = G^{(n)}(1)/n!$