

# STAT 3100 Lecture Note

Week Three (Sep 20 & 22, 2022)

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

2022/Sep/16 22:43:35

---

## Normal sampling theory (CB Sec. 5.3)

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

- If iid  $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ , then
  - $\sum_{i=1}^n X_i^2 \sim \chi^2(n)$  if iid  $X_1, \dots, X_n \sim \mathcal{N}(0, 1)$ ;
  - $X/\sqrt{Y/n} \sim t(n)$  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 (CB Def 5.5.20 & Thm 5.5.21)

### 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  $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$ ,

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

where  $f^{(n)}(x_0) = \frac{d^n}{dx^n} f(x)|_{x=x_0}$ .

- Called the Maclaurin series if  $x_0 = 0$

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

Under regularity conditions,

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

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

## Application

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

## Generating functions

### Moment generating function (mgf, CB Sec. 2.3)

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

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

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

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

- \*  $X_1, \dots, X_p$  are independent  $\Rightarrow M_{\mathbf{X}}(\mathbf{t}) = \prod_{i=1}^p M_{X_i}(t_i)$
- $M_{\mathbf{A}\mathbf{X}+\mathbf{b}}(\mathbf{t}) = \exp(\mathbf{b}^\top \mathbf{t}) M_{\mathbf{X}}(\mathbf{A}^\top \mathbf{t})$

- Application
  - Characterizing distributions:  $M_{\mathbf{X}}(\mathbf{t})$  and  $M_{\mathbf{Y}}(\mathbf{t})$  are both well-defined and equal for all  $\mathbf{t}$  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  $\mu'_n = EX^n = \sum_{k=0}^n \binom{n}{k} \mu_k (\mu'_1)^{n-k}$
    - \*  $n$ th central moment  $\mu_n = E(X - EX)^n = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} \mu'_k (\mu'_1)^{n-k}$

### Example Lec6.1

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

### Characteristic function

- For univariate  $X$ :  $\phi_X(t) = E\exp(itX)$  for all  $t \in \mathbb{R}$ 
  - Fourier transform of  $f_X$
  - Inverse:  $f_X(x) = (2\pi)^{-1} \int_{\mathbb{R}} \phi_X(t) \exp(-itx) dt$
  - $\mu'_n = EX^n = (-i)^n \phi_X^{(n)}(0)$

- For Multivariate  $\mathbf{X} = (X_1, \dots, X_p)^\top$ :  $\phi_{\mathbf{X}}(\mathbf{t}) = \mathbb{E} \exp(i\mathbf{t}^\top \mathbf{X})$  for all  $\mathbf{t} \in \mathbb{R}^p$ 
  - Fourier transform of  $f_{\mathbf{X}}$
  - Inverse:  $f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-p} \int_{\mathbb{R}^p} \phi_{\mathbf{X}}(\mathbf{t}) \exp(-i\mathbf{t}^\top \mathbf{x}) d\mathbf{t}$
- $\phi_{\mathbf{X}}(\mathbf{t}) = \phi_{\mathbf{Y}}(\mathbf{t})$  for all  $\mathbf{t} \in \mathbb{R}^p \Leftrightarrow \mathbf{X} \stackrel{d}{=} \mathbf{Y}$

### Example Lec6.2

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

### Other generating functions

- 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  $\{0, 1, \dots\}$ ,  $G(z) = \mathbb{E} z^X = \sum_{x=0}^{\infty} z^x p_X(x)$ .
  - $p_X(n) = \Pr(X = n) = G^{(n)}(1)/n!$