What to do today (Oct 14, 2020)?

2. Probability and Distribution (Chp 1-3)

2.4 Some Important Distributions (Chp3)

- *2.4.1 Discrete Distributions*
- *2.4.2 Continuous Distributions*
- *2.4.3 Multivariate Distributions*
- **2.4.4 Distributions Induced from Others**

3. Essential Topics in Mathematical Statistics (Chp 4-6)

3.1 Elementary Statistical Inferences (Chp 3)

- **3.1.1 Sampling and Statistics**
- *3.1.2 Confidence Interval*
- *3.1.3 Order Statistics*
- *3.1.4 Hypothesis Testing*
- *3.1.5 Statistical Simulation and Bootstrap*

2.4.4 Distributions Induced from Others: Distribution of Functions of Random Variables

rvs $X_1, \dots, X_K$, and their functions $Y_1 = g_1(X_1, \dots, X_K), \dots, Y_K = g_K(X_1, \dots, X_K)$. What are the distributions of $Y_1, \dots, Y_K$ if the joint distn of $X_1, \dots, X_K$ is known?

Suppose the mapping $\mathbf{g}$: $\mathbf{X} = (X_1, \dots, X_K) \to \mathbf{Y} = (Y_1, \dots, Y_K)$ is one-to-one. Thus

$$\mathbf{h} = \mathbf{g}^{-1}: \ \mathbf{Y} = (Y_1, \dots, Y_K) \to \mathbf{X} = (X_1, \dots, X_K).$$

That is, $X_1 = h_1(Y_1, \dots, Y_K), \dots, X_K = h_K(Y_1, \dots, Y_K)$.

Consider $K = 2$: in principle,

$$\begin{aligned} F_{\mathbf{Y}}(y_1, y_2) &= P(Y_1 \le y_1, Y_2 \le y_2) \\ &= P\big(g_1(X_1, X_2) \le y_1, g_2(X_1, X_2) \le y_2\big) \\ &= \int\int_{\substack{g_1(x_1, x_2) \le y_1 \\ g_2(x_1, x_2) \le y_2}} dF_{\mathbf{X}}(x_1, x_2). \end{aligned}$$

2.4.4 Distributions Induced from Others: Distribution of Functions of Random Variables

- if X_1, X_2 are discrete with the joint pmf $p_{\mathbf{X}}(x_1, x_2)$,

$$\begin{aligned} p_{\mathbf{Y}}(y_1, y_2) &= P\big(g_1(X_1, X_2) = y_1, g_2(X_1, X_2) = y_2\big) \\ &= P\big(X_1 = h_1(y_1, y_2), X_2 = h_2(y_1, y_2)\big) \\ &= p_{\mathbf{X}}\big(h_1(y_1, y_2), h_2(y_1, y_2)\big). \end{aligned}$$

- if X_1, X_2 are continuous with the joint pdf $f_{\mathbf{X}}(x_1, x_2)$,

$$f_{\mathbf{Y}}(y_1, y_2) = \frac{\partial^2 F_{\mathbf{Y}}(y_1, y_2)}{\partial y_1 \partial y_2}$$

$$= \begin{cases} \big|J\big| f_{\mathbf{X}}\big(h_1(y_1, y_2), h_2(y_1, y_2)\big), & (y_1, y_2) \in \text{the support of } (Y_1, Y_2) \\ 0, & \text{otherwise} \end{cases}$$

The Jacobian of the transformation is the determinant:

$$J = \left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \frac{\partial h_1(y_1, y_2)}{\partial y_1} & \frac{\partial h_1(y_1, y_2)}{\partial y_2} \\ \frac{\partial h_2(y_1, y_2)}{\partial y_1} & \frac{\partial h_2(y_1, y_2)}{\partial y_2} \end{vmatrix}$$

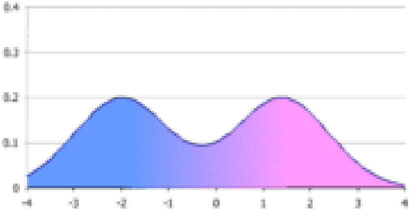
2.4.4 Distributions Induced from Others: Examples of Mixture Distributions

Consider two rvs X and W with joint distn $F(x, w)$.

- If the distn of X conditional on $W = w$ is $F(x|w)$, rv X's distn is $\int F(x|w)dF_W(w)$: a mixture of distributions $F(x|w)$ with all possible w's.
- **Example 1.** If W is discrete with pmf $p_W(w)$ for $w = w_1, \ldots, w_K$, X's distn is a mixture of the K distn's $F(x|w_1), \ldots, F(x|w_K)$:

$$X \sim F(x|w_1)p_W(w_1) + \ldots + F(x|w_K)p_W(w_K).$$

eg, a contaminated normal distn: if $W \sim B(1, p)$, $X = Z_1 W + Z_0(1 - W)$ with $Z_1 \sim N(\mu_1, \sigma_1^2), Z_0 \sim N(\mu_0, \sigma_0^2)$.



2.4.4 Distributions Induced from Others: Examples of Mixture Distributions

Consider two rvs X and W with joint distn $F(x, w)$.

- If the distn of X conditional on $W = w$ is $F(x|w)$, rv X's distn is $\int F(x|w)dF_W(w)$: a mixture of distributions $F(x|w)$ with all possible w's.
- **Example 2.** If W is continuous rv with pdf $f_W(w)$ for $w \in \mathcal{R}$, X's distn is a mixture of the distn's $\{F(x|w) : w \in \mathcal{R}\}$:

$$X \sim \int F(x|w)f_W(w)dw.$$

  - eg, Poisson-Gamma distn: when $X|W = \lambda \sim Poisson(\lambda)$ and $W \sim Gamma(\alpha, \beta)$.
  - Negative binomial distn is a Poisson-Gamma distn : $X \sim NB(r, \theta)$ if $W \sim Gamma(\alpha, \beta)$ with $\alpha = r$, an integer, and $\beta = (1 - \theta)/\theta$.

3.1 Elementary Statistical Inferences

3.1.1 Sampling and Statistics

In the information age, statistics are everywhere, since

- data are everywhere, and, on the other hand
- always resources are limited and our observation abilities are limited.

Various statistical methods.

- *to efficiently collect meaningful and sufficient information*: **Survey Sampling and Experimental Design**
- *to process the available information by tabulating/plotting the data*: **Descriptive Analysis**
- *to make inference about the target population, beyond what the information is directly on*: **Inferential Analysis**

Plus **Probability and Distribution**: *inferential reasoning with probability theory*

3.1.1 Sampling and Statistics

Why/What do we care about a random variable X? Its **distribution**: its pattern of taking different values, that is, what values X takes and how often it takes a particular value.

How do we find out X's distribution? From its observations (**data**: $x_1, \ldots, x_n$),

- (i) by picturing the data,
- (ii) by summarizing the data, and
- (iii) by estimating θ if $X \sim F(\cdot; \theta)$

Quesitons to ask:

- What to look for from the plot?
- Does $\bar{x}$ in general estimate μ well? What do we mean by "well"? How to assess a parameter estimator?
- Is there any systematic way to obtain a "good" estimator?
- What kind of observations are desirable?
- ... ...

What will we do in the next class?

1. *Introduction*

2. *Probability and Distribution (Chp 1-3)*

3. **Essential Topics in Mathematical Statistics (Chp 4-6)**
   - **3.1 Elementary Statistical Inferences**
     - **3.1.1 Sampling and Statistics**
     - **3.1.2 Confidence Interval**
     - **3.1.3 Order Statistics**
     - *3.1.4 Hypothesis Testing*
     - *3.1.5 Statistical Simulation and Bootstrap*
   - *3.2 Consistency and Limiting Distributions*
   - *3.3 Maximum Likelihood Methods*

4. *Further Topics, Selected from Chp 7-11*