- 추론 Inference
  - 추정 Estimation
    - 점추정 Point estimation
    - 구간추정 interval estimation
  - 검정 Test

- **Point estimation**
  - $X_i, \dots, X_n \overset{iid}{\sim} (\theta)$; random sample, n : sample size
  - $T = T(x_1, \dots, x_n)$: $X_1, \dots, X_n$의 조합으로 이루어진 통계량 statistic, 확률변수
  - X : 확률변수, x : 관측값 (관찰값)
  - $\theta \in \Omega$, Ω : 모수공간
  - 추정량 estimator

    T : θ의 추정량, 추정량 $\in$ 통계량

    $t = T(x_1, \dots, x_n)$: 추정치 estimate
  - 좋은 추정량의 조건

    unbiasedness : $E(T) - \theta = 0$
    - T는 θ의 불편추정량이다.

- **Maximum likelihood estimator**
  - $X_1, \dots, X_n \overset{iid}{\sim} f(x;\theta)$ $\Rightarrow$ 동일한 pdf $f(x;\theta)$ 를 갖는 분포로부터 뽑힌 독립인 n개의 표본
  - 목표 : θ 추정 - $X_1, \dots, X_n$을 사용

    $X_1, \dots, X_n$의 joint pdf

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

    θ를 모수로 갖을때 $(X_1 = x_1, \dots, X_n = x_n)$ 으로 관측될 가능성

    Change the role of X and θ

$$L(\theta; x_1, \dots, x_n) \text{ 가능도함수, } f(x_1, \dots, x_n; \theta) \text{의미 동일}$$

    IDEA : 가능도함수 $L(\theta; x_1, \dots, x_n)$를 최대로 해주는 θ를 θ의 추정량으로 선택

    $\Rightarrow \hat{\theta}^{mle} = \underset{\theta \in \Omega}{argmax} L(\theta; x_1, \dots, x_n) = \underset{\theta \in \Omega}{argmax} \ell(\theta; x)$

$$\begin{aligned} L(\theta; x_1, \dots, x_n) &= \prod_{i=1}^{n} f(x_i; \theta) \\ \ell(\theta; x_1, \dots, x_n) &= \log L(\theta; x_1, \dots, x_n) \\ &= \sum_{i=1}^{n} \log f(x_i; \theta) \end{aligned}$$

$\ell(\theta)$ 는 대부분 오목함수$^{\text{concave}}$

$$\frac{\partial \ell(\hat{\theta}^{mle})}{\partial \theta} = 0 \quad \Rightarrow \quad Estimating\ Equation$$

ex) Nomal $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, Find $\hat{\mu}^{mle}$, $\hat{\sigma^2}^{mle}$.

$$\begin{aligned} L(\theta; x_1, \ldots, x_n) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\} \\ &= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\} \\ \ell(\theta; x_1, \ldots, x_n) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} \end{aligned}$$

①

$$\begin{aligned} \frac{\partial \ell}{\partial \mu} &= \frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} \overset{set}{=} 0 \\ &\Leftrightarrow \sum_{i=1}^{n}(x_i-\mu) = 0 \\ &\Leftrightarrow \sum x_i = n\mu \\ &\therefore\ \mu = \bar{x},\ \hat{\mu}^{mle} = \bar{x} \end{aligned}$$

②

$$\begin{aligned} \frac{\partial \ell}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}\frac{1}{\sigma^4} \overset{set}{=} 0 \\ &\Leftrightarrow \frac{-n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4} = 0 \\ &\Leftrightarrow n\sigma^2 = \sum_{i=1}^{n}(x_i-\mu)^2 = 0 \\ &\therefore\ \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \\ &\therefore\ \hat{\sigma^2}^{mle} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \end{aligned}$$

ex) Uniform $X_1, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$, Find $\hat{\theta}^{mle}$

$$
\begin{aligned}
L(\theta; x_1, \ldots, x_n) &= \prod_{i=1}^{n} \frac{1}{\theta}\ I_{(0 \le x_i \le \theta)} \\
&= \frac{1}{\theta^n}\ I_{(min(x_i) \ge 0)}\ I_{(max(x_i) \le 0)}
\end{aligned}
$$

$\Rightarrow$ $\frac{1}{\theta^n}$: θ 에 대한 감소함수 (θ가 가장 작을때 , $\frac{1}{\theta^n}$ 은 가장 커진다)

$$\therefore\ \theta = max(x_i), \quad \hat{\theta}^{mle} = max(x_i)$$

- **Confidence Intervals** : θ가 속해있을 것이라 생각되는 구간을 제시

$$P(\theta \in [L, U]) = 1 - \alpha, \quad \alpha : \text{fixed value}$$

  - $\alpha = 0$인 경우는 의미 없음 $\Rightarrow$ 아무런 정보를 주지 않음
  - α를 적당히 작은 값으로 고정
  - 구간 길이가 가장 작을때 좋은 구간추정량이다.

  $\star$ **구간추정, 검정을 어떻게 할것인가?**

  Pivot random variable (Pivotal quaritity)

  $\Rightarrow$ 추정 대상 모수 θ와 θ의 좋은 점추정량으로 이루어진 **잘 알려진 분포**를 가지고 있는 확률변수 $\Rightarrow$ 확률 계산을 하기 위해 잘 알려진 분포를 따르는 확률변수를 사용한다.

- **Order statistics**

  $X_1, \ldots, X_n \overset{iid}{\sim} f(x;\theta) \ \Rightarrow\ min(X_i) = Y_1 < Y_2 < \cdots < Y_n = max(X_i)$ 으로 재배열

  $Y_k : X_1, \ldots, X_n$ 중에서 크기가 k번째로 작은 값

  ① $Y_1, \ldots, Y_n$ 의 joint pdf $\Rightarrow$ 변수변환

$$g(y_1, \ldots, y_n) = \begin{cases} n!f(y_1)\cdots f(y_n) & a < y_1 < \cdots < y_n < b \\ 0 & elsewhere \end{cases}$$

  X를 나열하는 모든 경우의 수만큼의 항이 나오게 된다. 각 항은 $f(y_1)\cdots f(y_2)|J|$ 이다. 여기서 모든 항의 $|J|$ 은 1이다. 따라서 나열하는 모든 경우의 수 ($n!$) 만큼 $f(y_1)\cdots f(y_2)$의 합이 된다.

  ② Population median

$$F(x) = P(X \le x) = \int_a^x f(w)dw, \quad a < x < b$$

  $\Rightarrow$ $F(m) = \frac{1}{2}, \quad m$: population median ex) $n = 3$

$X_1, X_2, X_3 \overset{iid}{\sim} f(x) \Rightarrow Y_1 < Y_2 < Y_3, \quad Y_2 :$ sample median

(1) Y_1, Y_2, Y_3 의 joint pdf

$$g(y_1, y_2, y_3) = \begin{cases} 6f(y_1)f(y_2)f(y_3) & a < y_1 < y_2 < y_3 < b \\ 0 & elsewhere \end{cases}$$

(2) Y_2 의 marginal pdf

$$\begin{aligned} h(y_2) &= \int_{y_2}^{b} \int_{a}^{y_2} 6f(y_1)f(y_2)f(y_3)dy_1dy_3 \\ &= \int_{y_2}^{b} 6f(y_2)f(y_3) \int_{a}^{y_2} f(y_1)dy_1dy_3 \\ &= 6f(y_2)F(y_2) \int_{y_2}^{b} f(y_3)dy_3 \\ &= \begin{cases} 6f(y_2)F(y_2)(1-F(y_2)), & a < y_2 < b \\ 0, & elsewhere \end{cases} \end{aligned}$$

(3) Compute $F(m) = P(Y_2 \leq m)$

$$\begin{aligned} &= \int_{a}^{m} 6f(y_2)F(y_2)(1-F(y_2))dy_2 \\ &= \int_{a}^{m} 6f(y_2)F(y_2) - 6f(y_2)F(y_2)^2 dy_2 \\ &= \left[6\frac{F(y_2)^2}{2} - 6\frac{F(y_2)^3}{3} \right]_a^m \\ &= 6\left[\frac{F(m)^2}{2} - 6\frac{F(m)^3}{3} \right] \\ &= 3\frac{1}{4} - 2\frac{1}{8} = \frac{1}{2} \\ &= F_{y_2}(m) \end{aligned}$$

$\Rightarrow$ 표본중위수 (Y_2) 의 중위수가 모집단의 중위수 (m) 가 된다.

③ Marginal pdf

ex) $n = 3$

$$\begin{aligned} h(y_2) &= \int_{y_2}^{b}\int_{a}^{y_2} 6f(y_1)f(y_2)f(y_3)dy_1dy_3 \\ &= \int_{y_2}^{b} 6f(y_2)f(y_3)\int_{a}^{y_2} f(y_1)dy_1dy_3 \\ &= 6f(y_2)F(y_2)\int_{y_2}^{b} f(y_3)dy_3 \\ &= \begin{cases} 6f(y_2)F(y_2)(1-F(y_2)), & a<y_2<b \\ 0, & elsewhere \end{cases} \end{aligned}$$

ex) $n = 4$

$$h(y_3) = \int\int\int 4!f(y_1)f(y_2)f(y_3)f(y_4)dydydy$$

(Generalization)

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!1!}f(y_k)[F(y_k)]^{k-1}[1-F(y_k)]^{n-k}$$

④ Heuristic derivation

$$P\left(x - \frac{\Delta}{2} \le x \le x + \frac{\Delta}{2}\right) \approx \Delta f(x) \ \ (\text{밑변}) \times (\text{높이})$$

⑤ Joint pdf of any two order statistics ($Y_i < Y_j$)

$$g_{ij}(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(y_i)]^{i-1}[F(y_j)-F(y_i)]^{j-i-1}[1-F(y_j)]^{n-j}f(y_i)f(y_j)$$