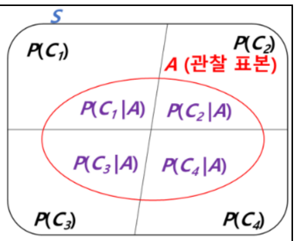
1. 확률, 확률분포

조건부 확률 / 독립성



표본공간 (Sample space): 가능한 실험결과 집합 ex) 주사위 $S = \{1,2,3,4,5,6\}$

확률: 특정 결과의 표본 공간내 점유도 ex) $P(X = 3) = 1/6$

1. **조건부**: $P(B|A) = P(A \cap B|A) = {P(A \cap B)}/{P(A)}$ ⇔ $P(A \cap B) = P(B|A)\,P(A)$

2. **Bayes** (C_i는 상호 배반=disjoint, S의 partition)

 1) **전확률 법칙**(LTP): $\boldsymbol{P(A) = \sum P(A \cap C_i) = \sum P(A \mid C_i)\, P(C_i)}$

 2) **베이즈 정리**: $\boldsymbol{P(C_i \mid A)} = {P(A \cap C_i)}/{P(A)} = {P(A \cap C_i)}/{\sum P(A \cap C_i)} = \boldsymbol{{P(A \mid C_i)\, P(C_i)}/{\sum P(A \mid C_i)\, P(C_i)}}$

  ① $P(C_i)$: C_i 사전확률 (prior)

  ② $P(C_i \mid A)$: C_i 사후확률 (posterior) ← 표본 A에서 관찰된 C_i

3) 독립성: $P(A \cap B \cap C) = P(A)P(B)P(C)$ 이면 A,B,C는 statistically independent

확률변수

확률변수: 표본공간의 특정 **사건**에 할당된 **수치값** ex) 베르누이 $S = \{H, T\}$ → $X \in \{0, 1\}$

1. Prob Mass Function; PMF (discrete) → CDF of PMF: $F(x) = P\big((-\infty, x]\big) = \sum_{(-\infty,x]} p(x)$

 ***변환**: $p_y(y) = p_X(w(y))$; 1-on-1 function $x = w(y)$

2. Prob Density Function; PDF (continuous) >0

 1) $F(x) = \int_{-\infty}^{x} f(t)dt$ ⇔ 2) $\frac{\mathrm{d}}{\mathrm{dx}}F(x) = f(x)$ (F는 f의 CDF)

 3) $P[(a,b)] = \int_{\mathrm{a}}^{\mathrm{b}} f(x)dx = F(b) - F(a)$

 ***변환**: X가 pdf f_X on S_X & Y가 pdf f_Y on S_Y; 1-on-1 $\boldsymbol{w(y) = x}$

 → $f_Y(y) = f_x\big(w(y)\big)|dx/dy|$ ⇔ $f_Y(y) = f_x\big(w(y)\big)\,\mathbf{abs}(\boldsymbol{J})$ (Jacobian: $J=|w'(y)|$ / $J=\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$) *J는 행렬식

* Support(받침): PDF $\neq 0$인 space // ***CDF는 유일** for PDF, PMF

기대값

1. 조건: $\boldsymbol{E(|X|)}$ **존재** ⇔ ① 연속 pdf 존재 ② $\int_{-\infty}^{\infty} \boldsymbol{|x| f(x)dx} < \infty$ (이산 pmf 존재 → $\sum |x_i|\, p_i(x) < \infty$)

2. 기대값: 1) <u>연속</u> $\boldsymbol{E(X) = \int_{-\infty}^{\infty} x\, f(x)dx}$ 2) <u>이산</u> $\boldsymbol{E(X) = \sum x_i\, p(x_i)}$

3. $y = g(x)$: 1) $E\big(g(x)\big) = \int_{-\infty}^{\infty} g(x)f(x)dx$ & $E\big(g(x)\big) = \sum g(x_i)p(x_i)$

 2) $E\big(k_1 g_1(x) + k_2 g_2(x)\big) = k_1 E\big(g_1(x)\big) + k_2 E\big(g_2(x)\big)$

1. 평균: $\mu = \boldsymbol{E(X)}$

2. 분산: $\sigma^2 = Var(X) = \boldsymbol{E[(X-\mu)^2]} = E(X^2) - \mu^2 = M''(0) - M'(0)^2$ *$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$

3. 적률생성함수 (MGF) *조건: $\mathrm{t} \in (-\mathrm{h}, \mathrm{h})$ for $\forall h > 0$ ← 0을 포함하는 개구간에서 mgf 존재

 1) $\boldsymbol{M(t) = E(e^{tX})}$ → $M_X(0)^{(r)} = E(X^r)$ *분포의 r차 moment

  ① $M(0)^{(r)} = \frac{d^r}{dt^r}\int_{-\infty}^{\infty} e^{tx}f(x)dx\,|_{t=0} = \int_{-\infty}^{\infty}\frac{d^r}{dt^r}e^{tx}f(x)dx\,|_{\mathrm{t=0}} = \int_{-\infty}^{\infty} x^r e^{tx}f(x)dx\,|_{\mathrm{t=0}} = \int_{-\infty}^{\infty} x^r\, f(x)dx = E(X^r)$

  ② $M(0)^{(r)} = \frac{d^r}{dt^r}\sum e^{tx_i}\, p(x_i)|_{\mathrm{t=0}} = \sum \frac{d^r}{dt^r} e^{tx_i}\, p(x_i) = \sum x_i^r e^{tx_i}\, p(x_i)|_{t=0} = \sum x_i^r\, p(x_i) = E(X^r)$

 2) 성질 ① MGF의 **유일성**: $\boldsymbol{M_x(t) = M_y(t)}$ ⇔ $\boldsymbol{X = Y}$ (pdf 동일)

  ② $M_{X+\alpha}(t) = e^{\alpha t}\, M_X(t)$ $\because M_{X+\alpha}(t) = E\big(e^{t(x+\alpha)}\big) = e^{\alpha t}\, E(e^{tx}) = e^{\alpha t}\, M_X(t)$

  ③ $M_{\alpha X}(t) = M_X(\alpha t)$ $\because M_{\alpha X}(t) = E\big(e^{t(\alpha X)}\big) = E\big(e^{(\alpha t)X}\big) = M_X(\alpha t)$

  ④ $M_{aX+b}(t) = e^{bt}M_X(at)$

  ⑤ $M_Y(t) = \prod M_{X_i}(k_i t)$, $t < |\min(h_i)|$ (for $Y = \sum k_i X_i$, X_i는 모두 <u>**독립**</u>)

  $\because M_Y(t) = E(\exp(tY)) = E(\exp(t\sum k_i X_i)) = E(\prod \exp(tk_i X_i)) = \prod E(\exp(tk_i X_i))$

  ⑥ $M_Y(t) = [M(t)^n]$ (for $Y = \sum X_i$; X_i는 **iid** 확률변수)

중요한 부등식

1. $E(X^m)$이 존재하면 → $E(X^k)$ 존재 for $\mathrm{k \leq m}$

 *증명: $E(X^k) = \int_{-\infty}^{\infty}|x|^k f(x)dx = \int_{|x|\leq 1}|x|^k f(x)dx + \int_{|x|\geq 1}|x|^k f(x)dx \leq \int_{|x|\leq 1} f(x)dx + \int_{|x|\geq 1}|x|^m f(x)dx$

 $\leq \int_{-\infty}^{\infty} f(x)dx + \int_{-\infty}^{\infty}|x|^m f(x)dx \leq 1 + E(X^m)$ ∴ 유한함

2. **Markov**: $\boldsymbol{P[u(X) \geq c] \leq E[u(X)]/c}$ (for u(X)≥0, c>0; E[u(X)]존재)

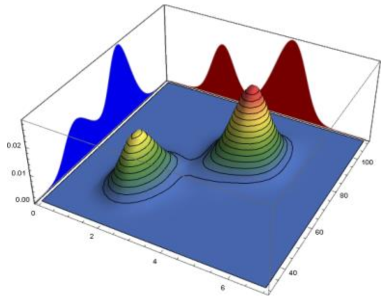
 *증명: $E[u(x)] = \int_{-\infty}^{\infty} u(x)f(x)dx \geq \int_{u(x)\geq c} u(x)f(x)dx \geq c\int_{u(x)\geq c} f(x)dx = c\,P[u(x) \geq c]$

 *직관: 평균 나이의 5배보다 나이 많은 사람의 확률 한계 → $P[X \geq 5\mu] \leq \frac{1}{5}$

3. **Chevyshev**: $\boldsymbol{P(|X-\mu| \geq k\sigma) \leq 1/k^2}$ (for k>0; X가 μ,σ^2(유한) 가짐)

 *증명: Markov에서 $u(X) = (X-\mu)^2$, $c = k^2\sigma^2$

2-1. 이변량분포

이변수	1) Joint CDF: $F(x,y)=P[\{X\le x\}\cap\{Y\le y\}]$ * $\mathbf{X}=(X,Y)^T\in D$; **Random vector X** *$P((a_1,a_2]\times(b_1,b_2])=F(a_2,b_2)-F(a_1,b_2)-F(a_2,b_1)+F(a_1,b_1) = \int_{b_1}^{b_2}\int_{a_1}^{a_2} f(x,y)\,dxdy$ 2) Joint PMF: $\sum_y\sum_x p(x,y)=1$ 3) Joint PDF: $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dydx$ $(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy=1)$ ⇔ $\partial^2(F)/\partial x\partial y = f(x,y)$  4) Marginal dist: **X**의 **효과**만 봄; **Y**는 **전체 범위 포괄** * $F_X(x)=P(\{X\le x\})=P(\{X\le x\}\cap\{-\infty<Y<\infty\})$ ①PMF of x: $F_X(x)=\sum_{(-\infty,x]}\{\sum_{y\in(-\infty,\infty)}p(x,y)\}$ ➔ $p_X(x)=\sum_{y\in(-\infty,\infty)}p(x,y)$ ②PDF of x: $F_x(x)=\int_{-\infty}^{x}\{\int_{-\infty}^{\infty}\boldsymbol{f(x,y)dy}\}\,dx$ ➔ $f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy$
이변수 기대값	*$E(g(X,Y))$ 존재 조건 ⇔ $E(\lvert g(X,Y)\rvert)<\infty$ 기대값: $E(g(X,Y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,f(x,y)\,dxdy$ (이산형: $E(g(X,Y))=\sum\sum g(x,y)p(x,y)$ 1) $E(k_1g_1+k_2g_2)=k_1E(g_1)+k_2E(g_2)$ 2) $E(\mathbf{X})=[E(X)\ E(Y)]^{\mathrm{T}}=[\int_{-\infty}^{\infty} x\, f_x(x)dx \ \ \int_{-\infty}^{\infty} y\, f_y(y)dy]^{\mathrm{T}}$ 3) $M(t_1,t_2)=E(\exp(t_1X+t_2Y))$ ➔ $\mathbf{t}=(t_1,t_2)^T$에 대해 $M(\mathbf{t})=E(\exp(\mathbf{t}^T\mathbf{X}))$ $E(X^kY^m)=\frac{\partial^{k+m}}{\partial t_1^k\,\partial t_2^m}M(0,0)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^ky^m\exp(t_1x+t_2y)\,f(x,y)\,dxdy$
이변수 변환	*변환 조건: 1) $\mathbf{X}=(X_1,X_2)$의 받침 S 2) S→T 사상하는 일대일대응: $y_1=u_1(x_1,x_2)$ & $y_2=u_2(x_1,x_2)$ 3) T→S 사상하는 위 대응 역: $x_1=w_1(y_1,y_2)$ & $x_2=w_2(y_1,y_2)$
	1. 이산형 변환: $p_{\mathbf{Y}}(y_1,y_2)=p_{\mathbf{X}}[w_1(y_1,y_2),w_2(y_1,y_2)]$ for $(y_1,y_2)\in T$ & 나머지 pmf 0 * X_1, X_2 ➔ Y로만 변환 시, dummy 변수를 하나 더 만들어 Y_2로 지정해주고 marginal Y dist를 구함
	2. 연속형 변환: $f_{\mathbf{Y}}(y_1,y_2)=f_{\mathbf{X}}[w_1(y_1,y_2),w_2(y_1,y_2)]\lvert\boldsymbol{J}\rvert$ for $(y_1,y_2)\in T$ & 나머지 pdf 0
	* MGF 이용 변환: $E(\exp(tY))=E(\exp(t(X_1+X_2))$ ➔ MGF 유일성으로 Y의 PMF/PDF 구함
조건부	1. 조건부 PMF: $p_{2\mid1}(x_2\mid x_1)=p(x_1,x_2)/p_1(x_1)$
	2. 조건부 PDF: $f_{2\mid1}(x_2\mid x_1)=f(x_1,x_2)/f_1(x_1)$ (Marginal $f_1(x_1)$는 스케일러) ***X**의 **효과**, **Y**는 **특정 값 고정** 1) $P(a<Y<b\mid X=x)=\int_a^b f_{Y\mid X}(y\mid x)\,dy$ & $P(c<X<d\mid Y=y)=\int_c^d f_{X\mid Y}(x\mid y)\,dx$ 2) $P(-\infty<Y<\infty\mid X=x)=\int_{-\infty}^{\infty} f_{Y\mid X}(y\mid x)\,dy=\int_{-\infty}^{\infty}\frac{f(x,y)}{f_X(x)}dy=\frac{1}{f_X(x)}\int_{-\infty}^{\infty}f(x,y)\,dy$=1 3) 조건부 기대값: $E[\boldsymbol{u(Y)}\mid x]=\int_{-\infty}^{\infty}u(y)\,f_{Y\mid X}(y\mid x)\,dy$ ➔ **x의 함수** ① 조건부 평균: $E(Y\mid x)=\int_{-\infty}^{\infty}y\,f_{Y\mid X}(y\mid x)\,dy$ ② 조건부 분산: $\mathrm{Var}(Y\mid x)=E(Y^2\mid x)-[E(Y\mid x)]^2$
	* 정리: μ_Y 추정 ← $E(Y\mid X)$이 Y보다 더 신뢰도 높음 (Rao & Blackwell) 1) $E[E(Y\mid X)]=E(Y)$ 2) $\mathrm{Var}(E(Y\mid X))\le\mathrm{Var}(Y)$ *Y 분산 유한 * 증명: $E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\,dydx=\int_{-\infty}^{\infty}[\int_{-\infty}^{\infty}yf_{Y\mid X}(y\mid x)dy]\,f_X(x)dx=\int_{-\infty}^{\infty}E(Y\mid X)\,f_X(x)dx=E(E(Y\mid X))$
공분산 / 상관 계수	1.공분산: $\mathbf{Cov}(\boldsymbol{X},\boldsymbol{Y})=\boldsymbol{E}[(\boldsymbol{X}-\boldsymbol{\mu_X})(\boldsymbol{Y}-\boldsymbol{\mu_Y})]=\boldsymbol{E}(\boldsymbol{XY})-\boldsymbol{E}(\boldsymbol{X})\boldsymbol{E}(\boldsymbol{Y})$ *독립이면 $\mathrm{Cov}(X,Y)=0\Leftrightarrow E(XY)=E(X)E(Y)$
	2.상관계수: $\boldsymbol{\rho}=\mathbf{Cov}(\boldsymbol{X},\boldsymbol{Y})/\boldsymbol{\sigma_X\sigma_Y}=\boldsymbol{\sigma_{XY}}/\boldsymbol{\sigma_X\sigma_Y}$ $(-1\le\rho\le1)$➔$y=a+bx\ (b>0)$에 ρ의 강도로 집중 $(0<\rho\le1)$
	3.선형조건부평균: $E(Y\mid X)=a+bx$ ➔ $E(Y\mid X)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(X-\mu_x)$ & $E(\mathrm{Var}(Y\mid X))=\sigma_y^2(1-\rho^2)$ *회귀분석 모회귀계수 $\beta=\rho(\sigma_y/\sigma_x)=\mathrm{Cov}(X,Y)/Var(X)$; *X,Y 분산 유한
선형결합 기대값/ 공분산	* 선형결합: $T=\sum_{i=1}^n a_iX_i$, $W=\sum_{i=1}^m b_iY_i$ 1) $E(\boldsymbol{T})=\sum_{i=1}^n \boldsymbol{a_i}E(\boldsymbol{X_i})$ (*$E[\lvert X_i\rvert]<\infty$) ∵ $E(X_1+X_2)=\int\int(x_1+x_2)F(x_1,x_2)dx_1dx_2=E(X_1)+E(X_2)$ 2) $\mathrm{Cov}(\boldsymbol{T},\boldsymbol{W})=\sum\sum\boldsymbol{a_ib_j}\,\mathrm{Cov}(\boldsymbol{X_i},\boldsymbol{Y_j})$ (*$E[X_i^2]<\infty$, $E[Y_{ij}^2]<\infty$) ① $Var(\boldsymbol{T})=\mathrm{Cov}(T,T)=\sum_{i=1}^n\boldsymbol{a_i^2}\,\mathrm{Var}(\boldsymbol{X_i})+2\sum_{i<j}\boldsymbol{a_ib_j}\,\mathrm{Cov}(\boldsymbol{X_i},\boldsymbol{X_j})$ (*$E[X_i^2]<\infty$) ② $Var(\boldsymbol{T})=\mathrm{Cov}(T,T)=\sum_{i=1}^n\boldsymbol{a_i^2}\,\mathrm{Var}(\boldsymbol{X_i})$ (*$X_1,\cdots,X_n$이 유한 분산, **독립**)

2-2. 다변량분포

독립	*정의: $\boldsymbol{f(x,y)=f_x(x)f_y(y)}$ ⇔ **X,Y는 독립** $[x\in(a,b)\,\&\,y\in(c,d)]$ (받침이 수평/수직선 box에 존재해야 함) 1. 조건부 증명: $f_y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx=\int_{-\infty}^{\infty} f_{y\mid x}(y\mid x)\,f_x(x)\,dx=f_{y\mid x}(y\mid x)\int_{-\infty}^{\infty} f_x(x)dx=f_{y\mid x}(y\mid x)$ 2. 동치류 1) $\boldsymbol{f(x,y)=f_x(x)\,f_y(y)}$ 2) $\boldsymbol{F(x,y)=F_x(x)\,F_y(y)}$ *증명: $\partial^2F/\partial x\partial y=f_x(x)f_y(y)$ 3) $\boldsymbol{P(a<X<b,\ c<Y<d)=P(a<X<b)\,P(c<Y<d)}$ *증명: $P(a<X<b,c<Y<d)=F(b,d)-F(a,d)-F(b,c)+F(a,c)=[F_x(b)-F_x(a)][F_y(d)-F_y(c)]$ 4) $\boldsymbol{E[u(X)v(Y)]=E[u(X)]\,E[v(Y)]}$ $\Rightarrow E(XY)=E(X)E(Y)\Leftrightarrow \mathrm{Cov}(X,Y)=0$ 5) $\boldsymbol{M(t_1,t_2)=M(t_1,0)\,M(0,t_2)}$ *증명: $M(t_1,t_2)=E(e^{t_1X+t_2Y})=E(e^{t_1X})E(e^{t_2Y})=M(t_1,0)\,M(0,t_2)$ * $\boldsymbol{M(t_1,0)}$는 **marginal $\boldsymbol{f_x}$의 MGF** ($\because M(t_1,0)=E(e^{t_1x})=\int_{-\infty}^{\infty}e^{t_1x}\left[\int_{-\infty}^{\infty}f(x,y)dy\right]dx=\int_{-\infty}^{\infty}e^{t_1x}f_x(x)dx$)
다변수	* $\mathbf{x}=(x_1,x_2,\cdots,x_p)^T=\left(X_1(c),X_2(c),\cdots,X_p(c)\right)^T$ for 확률 실험 c ∈ C 1. 결합 확률 함수들 1) Joint CDF: $F(\mathbf{x})=P[\{X_1\le x_1\}\cap\{X_2\le x_2\}\cap\cdots\cap\{X_p\le x_p\}]$ 2) Joint PMF $F(\mathbf{x})=\sum_{w_1\le x_1}\cdots\sum_{w_p\le x_p}p(w_1,\cdots,w_p)$ 3) Joint PDF: $F(\mathbf{x})=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_p}f(x_1,\cdots,x_p)\,dx_p\cdots dx_1$ $\left(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,\cdots,x_p)\,dx_p\cdots dx_1=1\right)$ ⇔ $\partial^p\{F(\mathbf{x})\}/\partial x_1\cdots\partial x_p=f(\mathbf{x})$
	2. Marginal/Conditional 1) $f_1(x_1)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,\cdots,x_p)\,dx_2\cdots dx_p$ ➔ $f_{2,\cdots,p\mid 1}(x_2,\cdots,x_p\mid x_1)=f(\mathbf{x})/f_1(x_1)$ 2) $f_{2,4,5}(x_2,x_4,x_5)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\mathbf{x})\,dx_1dx_3dx_6$ ← $\mathbf{x}=(x_1,x_2,x_3,x_4,x_5,x_6)^T$ ➔ $f_{1,3,6\mid 2.4.5}(x_1,x_3,x_6\mid x_2,x_4,x_5)=f(\mathbf{x})/f_{2,4,5}(x_2,x_4,x_5)$
	3. 기대값 1) $E(u(\mathbf{x}))=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}u(x_1,\cdots,x_p)\,dx_1\cdots dx_p$ (존재성: $^\exists E(\lvert u(\mathbf{x})\rvert)$) *이산: $E(u(\mathbf{x}))=\sum_{x_1}\cdots\sum_{x_p}u(x_1,\cdots,x_p)$ 2) $E(\sum k_iY_i)=\sum k_iE(Y_i)$ 3) $E[u(X_2,\cdots,X_p)\mid x_1]=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\boldsymbol{u(x_2,\cdots,x_p)}\,f_{2,\cdots,p\mid 1}(x_2,\cdots,x_p\mid x_1)\,dx_2\cdots dx_p$
	4. 독립: $E(\prod u_i(X_i))=\prod E(u_i(X_i))$ 등 동치류 (f,F,P,E,M) ***iid** (independent and identically distributed): 여러 확률 변수가 서로 독립 & 동일한 분포
	5. 변환: **X**의 받침 S에 대해, X⇔Y가 일대일이 되는 $S_1,\cdots,S_k$의 부분 공간 상 각각의 야코비안 J_i 정의 $g(\mathbf{y})=\sum_{i=1}^{k}\lvert\boldsymbol{J_i}\rvert\,f[w_{1i}(\mathbf{x}),\cdots,w_{pi}(\mathbf{x})]$
Random matrix	* **Random matrix** $\mathbf{W}=[W_{ij}]$, W_{ij} $(1\le i\le m,\ 1\le j\le n)$ 1. $E(\mathbf{W})=[E(W_{ij})]$ (일렬로 배열하여 mn x 1의 벡터로 생각) 1) $E[\mathbf{AW}+\mathbf{BV}]=\mathbf{A}\,E[\mathbf{W}]+\mathbf{B}\,E[\mathbf{V}]$ (A,B: k x m 상수 행렬, W,V: m x n 확률 행렬) 2) $E[\mathbf{AWB}]=\mathbf{A}\,E(\mathbf{W})\,\mathbf{B}$ (A: k x m, W: m x n, B: n x l) 2. 분산-공분산 행렬 (Variance-Covariance matrix) * $\mathbf{X}=(X_1,X_2,\cdots,X_p)^{\mathrm{T}}$; 모든 VCM는 양의 반정부호(psd) 1) 정의: $\mathrm{Cov}(\mathbf{X})=E[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^{\mathrm{T}}]=[\sigma_{jk}]$ $(\boldsymbol{\mu}=E(\mathbf{X}))$ ➔ $\sigma_j^2=\mathrm{Var}(X_j)$ & $\sigma_{jk}=\mathrm{Cov}(X_j,X_k)$ 2) 정리 ① $\mathrm{Cov}(\mathbf{X})=\mathrm{E}(\mathbf{XX}^{\mathrm{T}})-\boldsymbol{\mu\mu}^{\mathrm{T}}$ $(\sigma_i^2<\infty)$ ② $\mathrm{Cov}(\mathbf{AX})=\mathbf{A}\mathrm{Cov}(\mathbf{X})\mathbf{A}^{\mathrm{T}}$ $(\sigma_i^2<\infty,\ \mathrm{A{:}\ m\times p})$ $\because \mathrm{Cov}(\mathbf{AX})=\mathbf{A}\mathrm{E}(\mathbf{XX}^{\mathrm{T}})\mathbf{A}^{\mathrm{T}}-\mathbf{A}\mathrm{E}(\mathbf{X})\mathrm{E}(\mathbf{X})^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ 3. MGF: $M(\mathbf{t})=E[\exp(\mathbf{t}^T\mathbf{X})]=E[\prod_{i=1}^{p}\exp(t_iX_i)]$ (*$\mathrm{X_i}$ 독립➔ $M(\mathbf{t})=M(t_1,\cdots,0)\cdots M(0,\cdots,t_p)=\prod_{i=1}^{p}E[\exp(t_iX_i)]$ 1) $M_{\mathbf{Y}}(\mathbf{t})=\prod M_{\mathbf{X}_i}(\mathbf{t})$ ($\mathbf{Y}=\sum\mathbf{X}_i$, 각 $\mathbf{X}_i\in\mathbb{R}^n$은 독립) 2) $M_{\mathbf{Y}}(\mathbf{t})=\mathrm{e}^{\mathbf{b}^Tt}M_{\mathbf{X}}(\mathbf{A}^T\mathbf{t})$ ($\mathbf{Y}=\mathbf{AX}+\mathbf{b}$; $\mathbf{A}{:}\,m\times p$; $\mathbf{t}\in\mathbb{R}^m$; $\mathbf{b}\in\mathbb{R}^m$)

3-1. 주요 분포: 이산

이항 분포	* Bernoulli experiment: 성공/실패로 서로 배반인 확률 실험 * Bernoulli trial: 베르누이 실험을 독립적으로 반복 (성공 확률 p 동일) * Bernoulli distribution의 유도: X(성공)=1, X(실패)=0 ➔ PMF: $p(x) = p^x(1-p)^{1-x}$ *$\mu = p,\ \sigma^2 = p(1-p)$
	* **Binomial distribution (이항분포)**: n회 반복한 **베르누이** 시행에서 **성공한 총 횟수** 분포 1. PMF: $\boldsymbol{p(x) = \binom{n}{x} p^x (1-p)^{n-x} \sim b(n,p)}$ $(x = 0,1,\cdots,n)$ 2. MGF: $M(t) = \sum e^{tx}p(x) = [(1-p)+(pe^t)]^n$ $(t \in \mathbb{R})$ 3. 기대값 1) $\boldsymbol{\mu = np}$ *$\mu = M'(0) = n[(1-p)+pe^t]^{n-1}(pe^t)\vert_{t=0} = np$ 2) $\boldsymbol{\sigma^2 = np(1-p)}$ *$\sigma^2 = M''(0) - \left(M'(0)\right)^2 = n(n-1)p^2 + np - (np)^2 = np(1-p)$ 4. 가법성: $Y = \sum X_i,\ X_i \sim B(n_i, \boldsymbol{p})$ ➔ $Y \sim B(\sum n_i, \boldsymbol{p})$ (증명) $M_Y(t) = \prod M_{X_i}(t) = \prod[(1-p)+(pe^t)]^{n_i} = [(1-p)+(pe^t)]^{\sum n_i}$ 5. **성공 비율**(Y): $Y = X/n$ ➔ $E(Y) = p,\ \mathrm{Var}(Y) = p(1-p)/n$ 6. 근사 ① 푸아송 근사: **이항분포** $\boldsymbol{b(n,p)} \xrightarrow{D}$ **푸아송분포** $(\boldsymbol{\mu = np})$ (MGF 극한으로 분포수렴 증명← n↑, p↓) ② 정규 근사: $b(n,p) \approx N(np, np(1-p))$ ∗ 연속성 수정: $\boldsymbol{P(a \le X \le b) = P\left(a - \frac{1}{2} < X < b + \frac{1}{2}\right)}$, 후자가 더 정확 ∗ $p(x)$가 성공확률(= 평균) p인 Bernoulli분포 ↔ $X \sim B(1,p)$ → $Y = \sum_{i=1}^{20} X_i$ (iid) 에 대해 $p(y)$는 20회 시행 중 평균 $20p$회 성공하는 Bernoulli ↔ $Y \sim B(20,p)$
	* **Multinomial distribution (다항분포)** 1. PMF: $p(x_1,\cdots,x_{k-1}) = \frac{n!}{(x_1)!\cdots(x_k)!}(p_1)^{x_1}\cdots(p_k)^{x_k}$ ➔ $p_k = 1 - \sum_{i=1}^{k-1} p_i$ & $x_k = n - \sum_{i=1}^{k-1} x_i$ 2. MGF: $M(t_1,\cdots,t_{k-1}) = (p_1e^{t_1} + \cdots + p_{k-1}e^{t_{k-1}} + p_k)^n$
	* R codes 1) dbinom (k,n,p): P(X=k) 2) pbinom (k,n,p): P(X≤k)
	* **Negative binomial distribution (음이항분포)**: X번 실패 후 r번 성공 (베르누이 시행) *r번 성공시 나감 1. PMF: $p(x) = \binom{x+r-1}{r-1} p^r (1-p)^x$ 2. MGF: $M(t) = p^r[1-(1-p)e^t]^{-r}$ $(\mathrm{e^t} < 1/(1-\mathrm{p}))$ ⬄ [이항:x+(r-1)번 중 (r-1)번 성공] x [p] *$\binom{-n}{k} = (-n)(-n-1)\cdots(-n-k+1)/k! = (-1)^k\binom{n+k-1}{k}$ * **Geometric distribution (기하 분포)**: X번 실패 후 처음 성공 (베르누이 시행) ⬄ r=1인 음이항분포 1. PMF: $p(x) = p(1-p)^x$ 2. MGF: $M(t) = p[1-(1-p)e^t]^{-1}$
	* **Hypergeometric distribution (초기하분포)** 1. PMF: $p(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$ *N개 중 K개가 성공 & **비복원추출:** n번 시행 ➔ x번 성공 확률 2. 기대값: 1) $\mu = n\left(\frac{K}{N}\right)$ 2) $\sigma^2 = n\left(\frac{K}{N}\right)\left(1-\frac{K}{N}\right)\left(\frac{N-n}{N-1}\right)$ N>>n이면 이항분포로 근사 가능 $X \xrightarrow{D} B\left(n, \frac{K}{N}\right)$
Poisson 분포	* Poisson process: 일정한 구간 (**시간**, **공간**)에서 독립적으로 발생하는 event를 생성하는 과정 (**비기억성**) * Poisson postulate: 짧은 구간 h (h->0)에 대해 1) $g(1,h) = \lambda h + o(h)$ ∗ $g(x,w)$는 구간 길이 w 내에 x회 발생 확률 2) $\sum_{x=2}^{\infty} g(x,h) = o(h)$ (≒미소 구간 h에 둘 이상은 본질적 불가) * $\lim_{\mathrm{h}\to 0} o(h)/h = 0$ (little-o) 3) 겹치지 않는 구간 ➔ 확률적으로 독립 ∴ $g(x,w) = \frac{e^{-\lambda w}(\lambda w)^x}{x!}$
	1. PMF: $\boldsymbol{p(x)} = \frac{e^{-\mu}\mu^x}{x!}$ $(x = 0,1,2,\cdots)$ (*$\sum p(x) = e^{-\mu}\sum\left(\frac{\mu^x}{x!}\right) = e^{-\mu}e^{\mu} = 1$) ⟵ **주어진 시간/공간**에 **x회 발생** 분포 2. 기댓값: $\boldsymbol{\mu = \sigma^2 = \lambda w}$ (λ: 단위 길이당 발생률, w: 주어진 영역 크기) 3. MGF: $M(t) = e^{\mu(e^t - 1)}$ $(t \in \mathbb{R})$ 4. 가법성: $Y = \sum X_i,\ X_i \sim \mathrm{Poi}(m_i)$ ➔ $Y \sim \mathrm{Poi}(\sum m_i)$ (증명) $M_Y(t) = \prod M_{X_i}(t) = \prod e^{m_i(e^t-1)} = e^{(\sum m_i)(e^t-1)}$ ∗ $p(x)$가 주어진 100초당 평균 μ회 발생 Poisson ↔ $X \sim \mathrm{Poi}(\mu)$ → $Y = \sum_{i=1}^{20} X_i$ (iid) 에 대해 $p(y)$는 주어진 20×100초당 평균 20μ회 발생 Poisson ↔ $Y \sim \mathrm{Poi}(20\mu)$ * **이항분포** $\boldsymbol{b(n,p)} \xrightarrow{D}$ **푸아송분포** $(\boldsymbol{\mu = np})$
	* R codes 1) dpois (k,m): P(X=k) 2) ppois (k,m): P(X≤k)

3-2. 주요 분포: 감마 연관 분포

Γ 분포	* 감마함수: $\boldsymbol{\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}dy}$ $(\alpha > 0)$ * $\boldsymbol{\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)}$ ➔ $\boldsymbol{\Gamma(n) = (n-1)!}$ for 자연수 n *$\Gamma(1) = 1,\ \Gamma(1/2) = \sqrt{\pi}$ * 스털링 근사: $\Gamma(k+1) \approx \sqrt{2\pi k}\left(\frac{k}{e}\right)^k$
	***Gamma distribution (감마분포): α** (∈ ℝ) **번째** Poisson event 발생까지 걸리는 **대기 시간** 1. PDF: $\boldsymbol{f(x) = \frac{1}{\Gamma(k)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}}$ $\sim \boldsymbol{\Gamma(\alpha,\beta)}$ $(0 \le x < \infty)$ (감마함수 식에 $y = x/\beta$ 대입; α>0 & β>0) 2. 기댓값: 1) $\boldsymbol{\mu = \alpha\beta}$, 2) $\boldsymbol{\sigma^2 = \alpha\beta^2}$ 3. MGF: $M(t) = 1/(1-\beta t)^\alpha$ $(t < 1/\beta)$ 4. 가법성: $Y = \sum X_i,\ X_i \sim \Gamma(\alpha_i,\beta)$ ➔ $Y \sim \Gamma(\sum\alpha_i,\beta)$ (증명) $M_Y(t) = \prod M_{X_i}(t) = \prod(1-\beta t)^{-\alpha_i} = (1-\beta t)^{-\sum\alpha_i}$ 5. 스칼라배: $X \sim \Gamma(\alpha,\beta) \Rightarrow kX \sim \Gamma(\alpha,k\beta)$ (*증명: 야코비안 변수변환) 6. 유도: k번 Poisson event 발생까지 시간을 T_i로 분할 ➔ 각 $T_i \sim \Gamma(1,\frac{1}{\lambda})$ ➔ $Y = \sum_{i=1}^k T_i \sim \Gamma(k,\frac{1}{\lambda})$ (*Erlang 분포: 자연수 k인 감마 분포)
	* R codes 1) dgamma (x,shape=a,scale=b): f(X=x) 2) pgamma (x, shape=a, scale=b): P(X≤x)
지수 분포	***Exponential distribution (지수분포): 1번째** Poisson event 발생까지 **대기 시간** $= \Gamma(1,\frac{1}{\lambda})$ 1. PDF: $\boldsymbol{f(x) = \lambda e^{-\lambda x}}$ 2. 기댓값: 1) $\boldsymbol{\mu = 1/\lambda}$, 2) $\boldsymbol{\sigma^2 = 1/\lambda^2}$ 3. 유도: W가 첫 번째 Poisson event 까지 걸린 시간 ➔ w시간 내 푸아송 사건 없을 확률: $P(W > w) = \frac{e^{-\lambda w}(\lambda w)^0}{0!} = e^{-\lambda w}$ ⇔ $P(0 < W < w) = 1 - e^{-\lambda w}$ $\therefore f(w) = \lambda e^{-\lambda w}$
χ² 분포	***Chi-square distribution (카이제곱 분포)**: 자유도 r에 대해, $\boldsymbol{\chi^2(r) = \Gamma(\frac{r}{2},2)}$ 1. PDF: $\boldsymbol{f(x) = \frac{1}{\Gamma(\frac{r}{2})2^{(\frac{r}{2})}}x^{(\frac{r}{2})-1}e^{-\frac{x}{2}}}$ $\sim \boldsymbol{\chi^2(r)}$ $(0 \le x < \infty)$ 2. 기댓값: 1) $\boldsymbol{\mu = r}$, 2) $\boldsymbol{\sigma^2 = 2r}$ 3. MGF: $M(t) = 1/(1-2t)^{r/2}$ $(t < 1/2)$ 4. $E(X^k) = 2^k\frac{\Gamma(\frac{r}{2}+k)}{\Gamma(\frac{r}{2})}$, $k > -\frac{r}{2}$ 5. 가법성 (corollary): $Y = \sum X_i,\ X_i \sim \chi^2(r_i)$ ➔ $Y \sim \chi^2(\sum r_i)$
	* R codes 1) dchisq (x,r): f(X=x) 2) pchisq (x,r): P(X≤x)
β 분포	*베타함수: $B(\alpha,\beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy$ $(\alpha > 0, \beta > 0)$ ①$B(\alpha,\beta) = B(\beta,\alpha)$, ②$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$
	*결합 PDF: $h(x_1,x_2) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)}x_1^{\alpha-1}x_2^{\beta-1}e^{-(x_1+x_2)}$; $0 \le x_1 < \infty, 0 \le x_2 < \infty$ (X_1, X_2독립) * $Y_1 = X_1/(X_1+X_2)$ & $Y_2 = X_1 + X_2$ ➔ Y_1에 대한 marginal distribution이 beta(α,β) 1. PDF: $\boldsymbol{f(x) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}}$ $(0 < x < 1)$
	* R codes 1) dbeta(x,a,b): f(X=x) 2) pbeta (x,a,b): P(X≤x)
Dirichlet 분포 (β 확장)	*결합 PDF: $h(x_1,\cdots,x_{k+1}) = \prod_{i=1}^{k+1}\frac{1}{\Gamma(\alpha_i)}x_i^{\alpha_i-1}e^{-x_i}$; $0 \le x_i < \infty$ (X_1, … ,X_k 독립) * $Y_i = \frac{X_i}{\sum_{i=1}^{k+1}X_i}(i = 1,2,\cdots,k)$ & $Y_{k+1} = \sum_{i=1}^{k+1}X_i$ ➔$Y_1,\cdots,Y_k$에 대한 marginal dist이 Dirichlet(α_1,…α_{k+1}) 1. PDF: $g(\boldsymbol{y_1},\cdots,\boldsymbol{y_{k+1}}) = \boldsymbol{\frac{\Gamma(\alpha_1+\cdots+\alpha_{k+1})}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_{k+1})}y_i^{\alpha_i-1}\cdots y_k^{\alpha_k-1}[1-(y_1+\cdots+y_k)]^{\alpha_{k+1}-1}}$ $(0 \le y_k,\ \sum_{i=1}^k y_i < 1)$

3-3. 주요 분포: 정규 분포

정규 분포

*표준정규분포: $I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$ → $0 < \exp\left(-\frac{z^2}{2}\right) < \exp(-|z|+1)$ 유계 ($\int_{-\infty}^{\infty} e^{-|z|+1} dz = 2e$)

➔ $I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2+y^2}{2}\right) dxdy = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} r\, dr d\theta = 1$ $\quad \therefore f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$

*정규분포: $X = \sigma Z + \mu$ 로 변수 변환 $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$

*Bell shape 분포: location 모수 (μ), scale 모수 (σ^2) vs. 감마분포 등: shape 모수 (α), scale 모수 (β)

***표준 정규 분포 N(0, 1^2)**

1. PDF: $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right)$ $\quad (-\infty < z < \infty)$
2. MGF: $M(t) = \exp\left(\frac{1}{2} t^2\right)$, $t \in \mathbb{R}$ $\qquad$ 3. 기대값: $E(Z) = 0$, $\mathrm{Var}(Z) = 1$
4. $E(Z^k) = k! \Big/ 2^{\frac{k}{2}} \left(\frac{k}{2}\right)!$ (k가 짝수), $\quad E(Z^k) = 0$ (k가 홀수) $\qquad * M(t) = \exp\left(\frac{1}{2} t^2\right) = \sum_{m=0}^{\infty} \left(\frac{t^2}{2}\right)^m / m!$

***정규 분포 N(μ, σ^2)**

1. PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$ $\quad (-\infty < x < \infty,\ \sigma > 0)$
2. MGF: $M(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$, $t \in \mathbb{R}$ $\qquad$ 3. 기대값: $E(Z) = \mu$, $\mathrm{Var}(Z) = \sigma^2$
4. $E(X^k) = E\left[(\sigma Z + \mu)^k\right] = \sum_{j=0}^{k} \binom{k}{j} \sigma^j E(Z^j) \mu^{k-j}$
5. 가법성: $Y = \sum a_i X_i$, $X_i \sim N(\mu_i, \sigma_i^2)$ ➔ $Y \sim N[\sum(a_i\mu_i), \sum(a_i\sigma_i)^2]$

(증명) $M_Y(t) = \prod M_{a_i X_i}(t) = \prod M_{X_i}(a_i t) = \prod \exp\left(\mu_i(a_i t) + \frac{1}{2}\sigma_i^2 (a_i t)^2\right) = \exp\left((\sum a_i \mu_i)t + \frac{1}{2}(\sum a_i^2 \sigma_i^2)t^2\right)$

6. Corollary: $\bar{X} = \sum_{i=1}^{n} X_i / n$, $X_i \sim N(\mu, \sigma^2)$ (iid) ➔ $\bar{X} \sim N(\mu,\ \sigma^2/n)$

***표본 통계량 분포**

① $\bar{\boldsymbol{X}} \xrightarrow{D} \boldsymbol{N}\left(\mu, \frac{\sigma^2}{n}\right)$

② $\boldsymbol{S^2} \sim \sigma^2 \frac{\boldsymbol{\chi^2_{n-1}}}{n-1}$, $\quad E(S^2) = \sigma^2$ (CI: $\sigma^2 \in \left[\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right]$, $\mathrm{Var}(S^2) = \frac{2\sigma^4}{n-1}$

③ $\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F_{n-1,m-1}$ $\qquad \frac{\sigma_X^2}{\sigma_Y^2} \in \left[\frac{1}{F_{\frac{\alpha}{2},n-1,m-1}} \frac{s_x^2}{s_y^2},\ F_{\frac{\alpha}{2},m-1,n-1} \frac{s_x^2}{s_y^2}\right]$

*** 정리: $\boldsymbol{Z^2 \sim \chi^2(1)}$**

pf) $W = Z^2$일 때, $\quad F(x) = P(W \le x) = P(Z^2 \le x) = P(-\sqrt{x} \le Z \le \sqrt{x})$, $x \ge 0$

→ $y = \sqrt{w}$ 변환 시, $F(x) = 2\int_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = \int_0^x \frac{1}{\sqrt{2\pi}\sqrt{w}} \exp\left(-\frac{w}{2}\right) dw$

→ $f(x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}} = \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{1/2}} x^{\frac{1}{2}-1} e^{-\frac{x}{2}}$ $\sim \chi^2(1)$ $\quad (0 \le x < \infty)$

* **따름 정리**: $Y = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$ (가법성 of χ^2 using MGF; for iid Z ~ N(0,1^2))

---

* **Contaminated normal distribution**: 대부분 $Z \sim N(0,1^2)$, 일부 outlier $\sim N(0, \sigma_c^2)$ (오염 비율: ε)

1) $W = KZ + (1-K)\sigma_c Z$ $\quad$ for $K = \begin{cases} 1 & \text{확률 } 1-\varepsilon \\ 0 & \text{확률 } \varepsilon \end{cases}$ $\quad$ (Z, K는 독립)

$F_W(w) = P(W \le w) = P(W \le w, I = 1) + P(W \le w, I = 0) = P(Z \le w)(1-\varepsilon) + P\left(Z \le \frac{w}{\varepsilon_c}\right)\varepsilon = (1-\varepsilon)\Phi(w) + \varepsilon\,\Phi\left(\frac{w}{\sigma_c}\right)$

① PDF: $f_W(w) = (1-\varepsilon)\phi(w) + \frac{\varepsilon}{\sigma_c}\phi\left(\frac{w}{\sigma_c}\right)$ $\qquad$ ② $E(W) = 0$, $\mathrm{Var}(W) = 1 + \varepsilon(\sigma_c^2 - 1)$

2) $X = a + bW$ $(b > 0)$

① PDF: $f_X(x) = (1-\varepsilon)\phi\left(\frac{X-a}{b}\right) + \frac{\varepsilon}{\sigma_c}\phi\left(\frac{X-a}{b\sigma_c}\right)$ $\qquad$ ② $E(W) = a$, $\mathrm{Var}(W) = b^2[1 + \varepsilon(\sigma_c^2 - 1)]$

---

* R codes 1) dnorm(x,a,b): f(X=x) $\qquad$ 2) pnorm (x,a,b): P(X≤x) $\qquad$ (평균 a, 표준편차 b)

3-3. 주요 분포: 정규 분포

다변량 정규 분포

* $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I}_p)$ / $\mathbf{z} = (Z_1, \cdots, Z_p)^T \in \mathbb{R}^p$, $Z_j \sim$ iid $N(0,1)$

1) PDF: $f_{\mathbf{Z}}(\mathbf{z}) = \left(\frac{1}{2\pi}\right)^{p/2} \exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{z}\right)$ pf) $f_{\mathbf{Z}}(\mathbf{z}) = \prod \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z_j^2\right) = \left(\frac{1}{2\pi}\right)^{n/2} \exp\left(-\frac{1}{2}\sum z_j^2\right)$

2)MGF: $M_{\mathbf{Z}}(\mathbf{t}) = \exp\left(\frac{1}{2}\mathbf{t}^T\mathbf{t}\right)$ $(\mathbf{t} \in \mathbb{R}^{\boldsymbol{p}})$ pf) $M_{\mathbf{Z}}(\mathbf{t}) = E\{\exp(\mathbf{t}^T\mathbf{Z})\} = E\left\{\prod \exp(t_j Z_j)\right\} = \prod E\left\{\exp(t_j Z_j)\right\} = \exp\left(\frac{1}{2}\sum t_j^2\right)$

3) 기대값: $E[\mathbf{Z}] = \mathbf{0}$, $\mathrm{Cov}[\mathbf{Z}] = \mathbf{I}_p$

* $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ / $\mathrm{Cov}[\mathbf{X}] = \boldsymbol{\Sigma}$가 psd (양반정치)

⇔ p개의 **<u>의존관계</u>**인 정규분포 확률변수의 **결합 분포**

0) 변환: $\mathbf{X} = \boldsymbol{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}$ & $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$

1) PDF: $f_{\mathbf{X}}(\mathbf{x}) = \dfrac{1}{(2\pi)^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\dfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T(\boldsymbol{\Sigma}^{-1})(\mathbf{x} - \boldsymbol{\mu})\right\}$

2) MGF: $M_{\mathbf{X}}(\mathbf{t}) = \exp\left\{\mathbf{t}^T\boldsymbol{\mu} + \dfrac{1}{2}\mathbf{t}^T(\boldsymbol{\Sigma})\mathbf{t}\right\}$, $(\mathbf{t} \in \mathbb{R}^{\boldsymbol{p}})$

3) 기대값: $E[\mathbf{X}] = \boldsymbol{\mu}$, $\mathrm{Cov}[\mathbf{X}] = \boldsymbol{\Sigma}$

<유도> $\boldsymbol{\Sigma}$가 psd & 대칭 ➔ EVD 가능

$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}^T\boldsymbol{\Lambda}\boldsymbol{\Gamma}$ $(\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_p);$ $\lambda_1 \geq \cdots \geq \lambda_p)$

$\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Gamma}^T\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Gamma}$, $\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Gamma}^T\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\Gamma}$ (if $\boldsymbol{\Sigma}$ is pd)

$E[\mathbf{X}] = E[\boldsymbol{\Sigma}^{1/2}\mathbf{Z}] + \boldsymbol{\mu} = \boldsymbol{\Sigma}^{1/2}E[\mathbf{Z}] + \boldsymbol{\mu} = \boldsymbol{\mu}$

$\mathrm{Cov}[\mathbf{X}] = E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T] = E[(\boldsymbol{\Sigma}^{1/2}\boldsymbol{Z})(\boldsymbol{\Sigma}^{1/2}\boldsymbol{Z})^{\boldsymbol{T}}]$

$= \left(\boldsymbol{\Sigma}^{\frac{1}{2}}\right)E(\mathbf{Z}\mathbf{Z}^T)\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\right) = \boldsymbol{\Sigma}$ $* E[\mathbf{Z}\mathbf{Z}^T] = \mathrm{Cov}(\mathbf{Z}) + \mathbf{0} = \mathbf{I}_p$

1-1. Theorem $* \mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; $\mathbf{A}$: $\mathrm{m} \times \mathrm{p}$, $\mathbf{b} \in \mathbb{R}^{\mathrm{m}}$

$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ $\rightarrow \mathbf{Y} \sim N_{\mathrm{m}}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T)$ (MGF로 증명)

1-2. Corollary (m개 변수에 대한 **<u>주변 분포</u>**)

$*\mathbf{X} \rightarrow \mathbf{X_1} \in \mathbb{R}^m$, $\mathbf{X_2} \in \mathbb{R}^q$ $(\mathrm{p} = \mathrm{m} + \mathrm{q})$ 분할

- $\mathbf{X} = \begin{bmatrix}\mathbf{X_1}\\ \mathbf{X_2}\end{bmatrix}$, $\boldsymbol{\mu} = \begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix}$, $\boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}$
- $\mathbf{A} = [\mathbf{I}_m \quad \mathbf{O}_{mq}]$ $\rightarrow$ $\mathbf{X_1} = \mathbf{A}\mathbf{X}$

➔ $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \rightarrow \mathbf{X_1} \sim N_m(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$

$(\because A\mu = \mu_1$, $A\Sigma A^T = [\mathrm{I}_m \quad \mathrm{O}_{mq}]\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\begin{bmatrix}\mathrm{I}_m\\ \mathrm{O}_{mq}\end{bmatrix} = \Sigma_{11})$

<MGF 유도>

$M_{\mathbf{X}}(t) = \exp(\mathbf{t}^T\boldsymbol{\mu})\, M_{\mathbf{Z}}\{(\boldsymbol{\Sigma}^{\frac{1}{2}})^T\,\mathbf{t}\}$

$= \exp(\mathbf{t}^T\boldsymbol{\mu})\exp\{(1/2)[(\boldsymbol{\Sigma}^{1/2})^T\mathbf{t}]^{\boldsymbol{T}}[(\boldsymbol{\Sigma}^{1/2})^T\mathbf{t}]\}$

$= \exp(\mathbf{t}^T\boldsymbol{\mu})\exp[(1/2)\mathbf{t}^T(\boldsymbol{\Sigma}^{1/2})^T(\boldsymbol{\Sigma}^{1/2})\,\mathbf{t}] = e^{\mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T(\boldsymbol{\Sigma})\mathbf{t}}$

2. 주변분포 독립성: $\mathbf{X}_1, \mathbf{X}_2$ 독립 ⇔ $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \mathbf{0}$

pf) $M_{\mathbf{X}_1,\mathbf{X}_2}(\mathbf{t}_1, \mathbf{t}_2) = \exp\left\{[\mathbf{t_1} \quad \mathbf{t_2}]\begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix} + \frac{1}{2}[\mathbf{t_1} \quad \mathbf{t_2}]\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{t_1}\\ \mathbf{t_2}\end{bmatrix}\right\}$

$M_{\mathbf{X}_1}(\mathbf{t_1})M_{\mathbf{X}_2}(\mathbf{t_2}) = \exp\{\mathbf{t_1}\boldsymbol{\mu}_1 + \mathbf{t_2}\boldsymbol{\mu}_2 + \dfrac{1}{2}(\mathbf{t}_1^T\boldsymbol{\Sigma}_{11}\mathbf{t}_1 + \mathbf{t}_2^T\boldsymbol{\Sigma}_{22}\mathbf{t_2})\}$ $\therefore M_{\mathbf{X}_1,\mathbf{X}_2}(\mathbf{t}_1, \mathbf{t}_2) = M_{\mathbf{X}_1}(\mathbf{t_1})M_{\mathbf{X}_2}(\mathbf{t_2}) \leftrightarrow \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \mathbf{0}$

3. 조건부 분포: $\mathrm{X}_1|\mathrm{X}_2 \sim N_m(\boldsymbol{\mu_1} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X_2} - \boldsymbol{\mu_2}),\ \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$ ($\boldsymbol{\Sigma}$는 양정치)

pf) $\mathbf{W} = \mathbf{X_1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X_2}$ $\rightarrow$ $\begin{bmatrix}\mathbf{W}\\ \mathbf{X_2}\end{bmatrix} = \begin{bmatrix}\mathbf{I_m} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\\ \mathbf{O_{qm}} & \mathbf{I}_q\end{bmatrix}\begin{bmatrix}\mathbf{X_1}\\ \mathbf{X_2}\end{bmatrix}$ $\left(\mathbf{A} = \begin{bmatrix}\mathbf{I_m} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\\ \mathbf{O_{qm}} & \mathbf{I}_q\end{bmatrix}\right)$

$\begin{bmatrix}\mathbf{W}\\ \mathbf{X_2}\end{bmatrix} \sim N_p(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathbf{T}})$; $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathbf{T}} = \begin{bmatrix}\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{O_{mq}}\\ \mathbf{O_{qm}} & \boldsymbol{\Sigma}_{22}\end{bmatrix}$ ➔ **W, $\mathbf{X_2}$ 독립**

$\mathbf{W}|\mathbf{X_2} = \mathbf{W} \sim N_m(\boldsymbol{\mu_1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}) \rightarrow \therefore \boldsymbol{X_1}|\boldsymbol{X_2} \sim N_m(\boldsymbol{\mu_1} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X_2} - \boldsymbol{\mu_2}),\ \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$

4. 카이 제곱: $W = (\mathbf{X} - \boldsymbol{\mu})^T(\boldsymbol{\Sigma}^{-1})(\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Z}^T\mathbf{Z} \sim \chi^2(p)$ ($\boldsymbol{\Sigma}$는 양정치)

pf) $W = \mathbf{Z}^T\mathbf{Z} = \sum_{i=1}^{p} Z_i^2 \sim \chi^2(p)$ $*$ 가법성 of $\chi 2$ using MGF; for iid $\mathrm{Z} \sim \mathrm{N}(0,1^2)$ ➔ $\sum_{i=1}^{p}[(X_i - \mu_i)/\sigma_i]^2 \sim \chi^2(p)$

*** Bivariate normal distribution (이변량 정규 분포)**

1) **기댓값**: $\boldsymbol{\mu} = \begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix}$, $\boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\sigma}_1^2 & \boldsymbol{\sigma}_{12}\\ \boldsymbol{\sigma}_{12} & \boldsymbol{\sigma}_2^2\end{bmatrix}$, $\boldsymbol{\sigma}_{12} = \boldsymbol{\rho}\,\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2$

2) PDF: $f(x, y) = \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp\left\{-\dfrac{1}{2(1 - \rho^2)}\left[\left(\dfrac{x - \mu_1}{\sigma_1}\right)^2 + \left(\dfrac{y - \mu_2}{\sigma_2}\right)^2 - 2\rho\left(\dfrac{x - \mu_1}{\sigma_1}\right)\left(\dfrac{y - \mu_2}{\sigma_2}\right)\right]\right\}$

3) 조건부 분포: $Y|X \sim N[\mu_2 + \rho\dfrac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1 - \rho^2)]$

PCA 기본

$\mathbf{Y} = \boldsymbol{\Gamma}\mathbf{X} = (\mathbf{PC_1}, \mathbf{PC_2}, \cdots, \mathbf{PC}_p)^T$ ➔ $\mathbf{PC_1} = \mathbf{v}_1^T\mathbf{X}$ ($\mathbf{v_1}$: $\mathrm{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$의 λ_1 대응 고유벡터)

pf)$\mathbf{Y} \sim N_p(\boldsymbol{\Gamma}\boldsymbol{\mu}, \boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^T) = N_p(\boldsymbol{\Gamma}\boldsymbol{\mu}, \boldsymbol{\Lambda})$ $\rightarrow$ $\mathrm{TV}(\mathbf{X}) = \sum\sigma_i^2 = tr(\boldsymbol{\Sigma}) = tr(\boldsymbol{\Lambda}) = \sum\lambda_i = \mathrm{TV}(\mathbf{Y})$

어떤 $\|\mathbf{a}\|^2 = 1$, $\mathbf{a} = \sum_{j=1}^{p} a_j\mathbf{v}_j$ *에 대해* $\mathbf{a}^T\mathbf{v}_1 = a_i$

$\rightarrow \mathrm{Var}(\mathbf{a}^T\mathbf{X}) = \mathrm{Cov}(\mathbf{a}^T\mathbf{X}) = \mathbf{a}^T\Sigma\mathbf{a} = (\boldsymbol{\Gamma}\mathbf{a})^T\boldsymbol{\Lambda}(\boldsymbol{\Gamma}\mathbf{a}) = \sum_i^p\lambda_i(\mathbf{a}^T\mathbf{v}_i)^2 = \sum_i^p\lambda_i a_i^2 \leq \lambda_1 = \mathrm{Var}(Y_1)$, 등호조건: $\mathbf{a} = \mathbf{v}_1$

$\therefore Y_1 = \mathbf{v}_1^T\mathbf{X}$ (고유벡터 v_1으로 총 데이터 X 사영): 총분산 $\sum\lambda_i$ 중 최대 분산 λ_1 을 설명하는 $\mathbf{PC_1}$

$\rightarrow \mathbf{X} = \boldsymbol{\Gamma}^{\mathbf{T}}\mathbf{Y}$ 에서 $X_k = (v_{1k})\mathbf{PC_1} + (v_{2k})\mathbf{PC_2} + \cdots$ (각 v_{ik}는 X_k의 PC_i에 대한 PC score)

3-4. 표본통계량의 분포

표본 평균 $\bar{X}$

* $X_1, \cdots, X_n \overset{iid}{\sim} dist(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$

1) 표본평균: $\bar{\boldsymbol{X}} = \sum_{i=1}^{n} \boldsymbol{X_i} / \boldsymbol{n}$

① $\boldsymbol{E(\bar{X})} = \sum_{i=1}^{n} E(X_i)/n = n\mu/n = \boldsymbol{\mu}$

② $\mathbf{Var}(\bar{\boldsymbol{X}}) = \sum_{i=1}^{n} \left(\frac{1}{n}\right)^2 \mathrm{Var}(X_i) = n\left(\frac{1}{n}\right)^2 \sigma^2 = \boldsymbol{\sigma^2/n}$

$\therefore \bar{\boldsymbol{X}} \xrightarrow{D} \boldsymbol{N}\left(\boldsymbol{\mu}, \frac{\boldsymbol{\sigma^2}}{\boldsymbol{n}}\right)$ by CLT

표본 분산 S^2

2) 표본분산: $\boldsymbol{S^2} = \sum_{i=1}^{n}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^2 / \boldsymbol{n-1} = \sum X_i^2 - n\bar{X}^2 / n-1$ (σ^2의 불편추정량)

① $\boldsymbol{E(S^2)} = \sum E(X_i^2) - nE(\bar{X}^2) / n-1 = n(\mu^2+\sigma^2) - n(\mu^2 + \frac{\sigma^2}{n}) / n-1 = \boldsymbol{\sigma^2}$

② $\mathrm{Var}(S^2) = \frac{2\sigma^4}{n-1}$ ($\bar{X}$, S^2는 독립 by Student's 정리)

③ $\boldsymbol{S^2} \sim \sigma^2 \frac{\boldsymbol{\chi^2_{n-1}}}{n-1}$

표본 통계량 분포

① $\bar{\boldsymbol{X}} \xrightarrow{D} \boldsymbol{N}\left(\mu, \frac{\sigma^2}{n}\right)$

② $\boldsymbol{S^2} \sim \sigma^2 \frac{\boldsymbol{\chi^2_{n-1}}}{n-1}$, $E(S^2) = \sigma^2$ (CI: $\sigma^2 \in \left[\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right]$, $\mathrm{Var}(S^2) = \frac{2\sigma^4}{n-1}$

③ $\frac{s_X^2/\sigma_X^2}{s_Y^2/\sigma_Y^2} \sim F_{n-1,m-1}$ $\quad \frac{\sigma_X^2}{\sigma_Y^2} \in \left[\frac{1}{F_{\frac{\alpha}{2},n-1,m-1}}\frac{s_x^2}{s_y^2},\ F_{\frac{\alpha}{2},m-1,n-1}\frac{s_x^2}{s_y^2}\right]$

다변량 표본 분포

* $\{\mathbf{X}_i\} = [X_{i1}, X_{i2}, \cdots, X_{ij}, \cdots, X_{ip}]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{p}} \overset{iid}{\sim} dist(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ⇔ $X_{ij} \sim dist(\mu_j, \sigma_j^2)$

	모집단	표본
평균벡터	$\boldsymbol{\mu} = \begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_j\\ \vdots\\ \mu_p\end{bmatrix} = \begin{bmatrix}E(X_1)\\ E(X_2)\\ \vdots\\ E(X_j)\\ \vdots\\ E(X_p)\end{bmatrix} = E(\mathbf{X})$	$\bar{\mathbf{X}} = \begin{bmatrix}\bar{X}_1\\ \bar{X}_2\\ \vdots\\ \bar{X}_j\\ \vdots\\ \bar{X}_p\end{bmatrix} = \begin{bmatrix}\frac{1}{n}\sum_{i=1}^{n} X_{i1}\\ \frac{1}{n}\sum_{i=1}^{n} X_{i2}\\ \vdots\\ \frac{1}{n}\sum_{i=1}^{n} X_{ij}\\ \vdots\\ \frac{1}{n}\sum_{i=1}^{n} X_{ip}\end{bmatrix} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{\mathrm{i}}$
공분산행렬	$\boldsymbol{\Sigma} = \begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1j} & \cdots & \sigma_{1p}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2j} & \cdots & \sigma_{2p}\\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots\\ \sigma_{j1} & \sigma_{j2} & \cdots & \sigma_{jj} & \cdots & \sigma_{jp}\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pj} & \cdots & \sigma_{pp}\end{bmatrix}$ $= \mathrm{E}\left(\begin{bmatrix}X_1-\mu_1\\ \vdots\\ X_p-\mu_p\end{bmatrix}[X_1-\mu_1 \ \cdots \ X_p-\mu_p]\right)$ $= E\{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^{\mathrm{T}}\} = \mathrm{Cov}(\mathbf{X})$	$\mathbf{Q} = \begin{bmatrix}s_{11} & s_{12} & \cdots & s_{1j} & \cdots & s_{1p}\\ s_{21} & s_{22} & \cdots & s_{2j} & \cdots & s_{2p}\\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots\\ s_{j1} & s_{j2} & \cdots & s_{jj} & \cdots & s_{jp}\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ s_{p1} & s_{p2} & \cdots & s_{pj} & \cdots & s_{pp}\end{bmatrix}$ $= \frac{1}{n-1}\sum_{i=1}^{n}\begin{bmatrix}X_{i1}-\bar{X}_1\\ \vdots\\ X_{ip}-\bar{X}_p\end{bmatrix}[X_{i1}-\bar{X}_1 \ \cdots \ X_{ip}-\bar{X}_p]$ $= \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i-\bar{\mathbf{X}})(\mathbf{X}_i-\bar{\mathbf{X}})^T$
상관행렬	$\boldsymbol{\rho} = \begin{bmatrix}1 & \rho_{12} & \cdots & \rho_{1j} & \cdots & \rho_{1p}\\ \rho_{21} & 1 & \cdots & \rho_{2j} & \cdots & \rho_{2p}\\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots\\ \rho_{j1} & \rho_{j2} & \cdots & 1 & \cdots & \rho_{jp}\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pj} & \cdots & 1\end{bmatrix}$ $\rho_{jk} = \frac{\sigma_{jk}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}}$	$\mathbf{R} = \begin{bmatrix}1 & r_{12} & \cdots & r_{1j} & \cdots & r_{1p}\\ r_{21} & 1 & \cdots & r_{2j} & \cdots & r_{2p}\\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots\\ r_{j1} & r_{j2} & \cdots & 1 & \cdots & r_{jp}\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ r_{p1} & r_{p2} & \cdots & r_{pj} & \cdots & 1\end{bmatrix}$ $r_{jk} = \frac{s_{jk}}{\sqrt{s_{ii}}\sqrt{s_{kk}}}$

$\mathbf{X} \sim dist(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \ \Rightarrow \ E(\mathbf{X}^T\mathbf{A}\mathbf{X}) = tr(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}$ (A는 대칭행렬)

$\because E = E(\mathbf{X}^T\mathbf{A}\mathbf{X}) = E\left(tr(\mathbf{X}^T\mathbf{A}\mathbf{X})\right) = E\left(tr(\mathbf{A}\mathbf{X}\mathbf{X}^T)\right) = tr\left(\mathbf{A}E(\mathbf{X}\mathbf{X}^T)\right) = tr\left(\mathbf{A}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T)\right)$ (*tr은 선형연산자)

$\therefore X_i \overset{iid}{\sim} dist(\mu, \sigma^2)$ & $\mathbf{X} = [X_1, \cdots, X_n]^T \Rightarrow (n-1)S^2 = \mathbf{X}^T\left(\mathbf{I} - \frac{1}{\mathrm{n}}\mathbf{J}\right)\mathbf{X}$, $E(S^2) = \sigma^2$

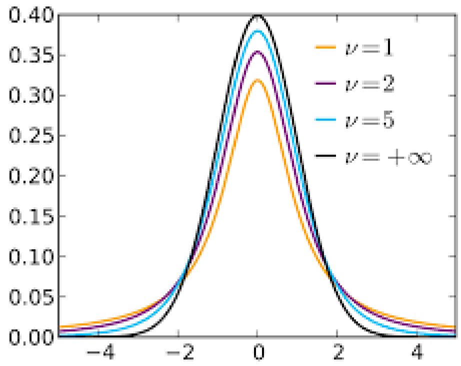
3-5. 주요 분포: t-분포, F-분포

t-분포

Joint PDF: $h(w,v)=\left\{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}w^2\right)\right\}\left\{\frac{1}{\Gamma\left(\frac{r}{2}\right)2^{\left(\frac{r}{2}\right)}}v^{\left(\frac{r}{2}\right)-1}e^{-\frac{v}{2}}\right\}$ $\quad w\in(-\infty,\infty), v\in(0,\infty)$; 서로 독립

$T=\frac{W}{\sqrt{V/r}}$ 일때 $t=\frac{w}{\sqrt{v/r}}, u=v$로 변환 시 $|J|=\frac{\sqrt{u}}{\sqrt{r}}$ $\to g(t,u)$ 에서 T의 marginal PDF $g_1(w)=\int_0^\infty g(t,u)du$

$T=\frac{W}{\sqrt{V/r}}\sim\frac{\boldsymbol{N(0,1)}}{\sqrt{\boldsymbol{\chi_r^2/r}}}$ $\leftarrow W\sim N(0,1^2), V\sim\chi^2(r)$; 서로 독립



1. PDF: $f(x)=\frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\,\Gamma\left(\frac{r}{2}\right)}\left(1+\frac{x^2}{r}\right)^{-\left(\frac{r+1}{2}\right)}$, $x\in\mathbb{R}$

2. 기대값: $E(T)=0\ (r>1)$, $\mathrm{Var}(T)=\frac{r}{r-2}\ (r>2)$

$*\ E(T^k)=E\left[W^k\left(\frac{V}{r}\right)^{-k/2}\right]=E(W^k)E\left[\left(\frac{V}{r}\right)^{-k/2}\right]=\frac{1}{r^{-\frac{k}{2}}}E(W^k)E\left(V^{-k/2}\right)$

$=E(W^k)\left(\frac{r}{2}\right)^{\frac{k}{2}}\frac{\Gamma\left(\frac{r}{2}-\frac{k}{2}\right)}{\Gamma\left(\frac{r}{2}\right)}$, $\quad r>k$

(Cauchy 분포: df = 1인 t 분포, $f(x)=\frac{1}{\pi(1+x^2)}$) $*$ df $\to\infty$ 이면 $\boldsymbol{T\xrightarrow{D}N(0,1)}$

3. Student's theorem *iid $X_i\sim N(\mu,\sigma^2)$ $\boldsymbol{n}$개 **[직관: $\boldsymbol{\sigma^2}$를 모르나, $\boldsymbol{X_i\sim N(\mu,\sigma^2)}$ ➔ $\boldsymbol{\bar{X}}$, $\boldsymbol{S}$로 $\boldsymbol{\mu}$ 정확 추정]**

1) $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$, $\quad$ 2) $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t\,(n-1)$ $\quad\leftarrow T=\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)\Big/\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}$

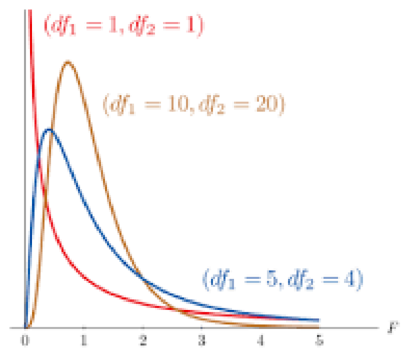
pf) $V=\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2=\sum_{i=1}^n\left(\frac{(X_i-\bar{X})+(\bar{X}-\mu)}{\sigma}\right)^2=\sum_{i=1}^n\left(\frac{X_i-\bar{X}}{\sigma}\right)^2+\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2=\frac{(n-1)S^2}{\sigma^2}+\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$

$\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$, $\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2\sim\chi^2(1)$ $\to\bar{X},S^2$ 독립; 양변 mgf 취하면 $(1-2t)^{-\frac{n-1}{2}}=E\left(\exp\left\{\frac{(n-1)S^2}{\sigma^2}\right\}t\right)$

* R codes 1) dt(x,r): f(X=x) 2) pt(x,r): P(X≤x) (자유도 r) // 3) qt(0.975,r): 97.5%인 t값

F-분포

$F=\frac{U/r_1}{V/r_2}\sim\frac{\boldsymbol{\chi_{r_1}^2/r_1}}{\boldsymbol{\chi_{r_2}^2/r_2}}$ $\leftarrow U\sim\chi^2(r_1), V\sim\chi^2(r_2)$; 서로 독립



1. PDF: $f(x)=\frac{1}{\mathrm{B}\left(\frac{r_1}{2},\frac{r_2}{2}\right)}\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}}\left(1+\frac{r_1}{r_2}x\right)^{-\frac{r_1+r_2}{2}}x^{\frac{r_1}{2}-1}$ $\quad 0<x<\infty$

2. 기댓값: $E(F)=\frac{r_2}{r_2-2}\quad(r_2>2)$

$*\ E(F^k)=E\left[\left(\frac{U/r_1}{V/r_2}\right)^k\right]=\left(\frac{r_2}{r_1}\right)^k E(U^k)E(V^{-k})$

$=\left(\frac{r_2}{r_1}\right)^k\frac{\Gamma\left(\frac{r_1}{2}+k\right)}{\Gamma\left(\frac{r_1}{2}\right)}\frac{\Gamma\left(\frac{r_2}{2}-k\right)}{\Gamma\left(\frac{r_2}{2}\right)}\quad(r_2>2k)$

* R codes 1) df(x,a,b): f(X=x) 2) pf(x,a,b): P(X≤x) (자유도 a,b) // 3) qf(0.975,a,b): 97.5%인 F값

혼합 분포

1. 혼합분포: $f(x)=\sum_{i=1}^k p_i f_i(x)$ (각각 받침 S_i, 평균 μ_i, 분산 σ_i^2 , $\sum_{j=1}^k p_j=1$)

1) $E(X)=\sum_{i=1}^k p_i\int_{-\infty}^{\infty}xf_i(x)dx=\sum_{i=1}^k p_i\mu_i=\bar{\mu}$

2) $Var(X)=\sum_{i=1}^k p_i\sigma_i^2+\sum_{i=1}^k p_i(\mu_i-\bar{\mu})^2$

$*\ Var(X)=\sum_{i=1}^k p_i\int_{-\infty}^{\infty}(x-\bar{\mu})^2 f_i(x)dx=\sum_{i=1}^k p_i\int_{-\infty}^{\infty}(x-\mu_i)^2 f_i(x)dx+\sum_{i=1}^k p_i(\mu_i-\bar{\mu})^2\int_{-\infty}^{\infty}f_i(x)dx$

2. 응용: 유한 개로 제한할 필요 X

1) **감마(θ,α,β)+푸아송(x,θ) ➔ 음이항 분포**

$p(x)=\int_0^\infty\left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\theta^{\alpha-1}e^{-\theta/\beta}\right]\left[\frac{e^{-\theta}\theta^x}{x!}\right]d\theta=\frac{1}{\Gamma(\alpha)\beta^\alpha x!}\int_0^\infty\theta^{\alpha+x-1}e^{-\theta(\alpha+\beta)/\beta}d\theta=\frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!}\frac{\beta^x}{(1+\beta)^{\alpha+x}}$

$\to p(x)=\frac{(r+x-1)!}{(r-1)!}\frac{p^r(1-p)^x}{x!}$, $\quad\alpha=r\in\mathbb{Z}^+$, $\quad\beta=\frac{1-p}{p}$

2) **베이지안 추론**: $h(x)=\int_\theta g(\theta)f(x|\theta)d\theta$; $g(\theta)$: Conjugate prior, $h(x)$: 무조건부

① $X|\theta\sim\mathrm{N}(0, 1/\theta)$, $\theta\sim\Gamma(r/2, 2/r)\to X\sim t(r)$ ② 이항분포 (p모름)→베타분포 β(p)로 추출 $\int_0^1 p(x|p)g(p)dp$

4. 일치성 / 극한분포 ("통계학적 수렴")

중요한 부등식

1. **Markov**: $\boldsymbol{P[u(X) \geq c] \leq E[u(X)]/c}$ (for u(X)≥0, c>0; E[u(X)]존재)

   *증명: $E[u(x)] = \int_{-\infty}^{\infty} u(x)f(x)dx \geq \int_{u(x)\geq c} u(x)f(x)dx \geq c\int_{u(x)\geq c} f(x)dx = c\,P[u(x) \geq c]$

2. **Chevyshev**: $\boldsymbol{P(|X-\mu| \geq k\sigma) \leq 1/k^2}$ (for k>0; X가 μ,σ²(유한) 가짐)

   *증명: Markov에서 $u(X) = (X-\mu)^2,\ c = k^2\sigma^2$

확률 수렴

1. **정의**: $\boldsymbol{X_n \xrightarrow{P} X}$ ⇔ $\forall\epsilon > 0,\ \boldsymbol{\lim_{n\to\infty} P[|X_n - X| \geq \epsilon] = 0}$ $\left(\Leftrightarrow \lim_{n\to\infty} P[|X_n - X| < \epsilon] = 1\right)$ "함수열의 수렴"

   ($\boldsymbol{X_n \xrightarrow{P} a}$, if X가 상수 a ➔ "퇴화확률변수, p(a)=1, 나머지 0")

2. **대수의 약법칙**: iid $\{X_n\} \sim \left(\textbf{평균}: \boldsymbol{\mu}, \textbf{분산}: \boldsymbol{\sigma^2 < \infty}\right)$, $\boldsymbol{\bar{X}_n \xrightarrow{P} \mu}$

   *증명: By Chevyshev's ineq, $\boldsymbol{P(|\bar{X}_n - \mu| \geq \epsilon) \leq \sigma^2/(n\epsilon^2)} \to 0$ (when $n \to \infty$)

3. **정리**

	정리	증명
선형	$* X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$ ① $\boldsymbol{(X_n + Y_n) \xrightarrow{P} (X+Y)}$ ② $\boldsymbol{aX_n \xrightarrow{P} aX}$	①P는 집합오염에 단조 (=공간 커지면 확률 커짐); 삼각부등식 $P[|(X_n+Y_n)-(X-Y)| \geq \epsilon] \leq P[|X_n - X| + |Y_n - Y| \geq \epsilon]$ $\leq P[|X_n - X| \geq \epsilon/2] + P[|Y_n - Y| \geq \epsilon/2]$
함수	* 받침 상 연속 $g(x)$ ③ $\boldsymbol{X_n \xrightarrow{P} a \Rightarrow g(X_n) \xrightarrow{P} g(a)}$ ④$\boldsymbol{X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X)}$	① $|g(x)-g(a)| \geq \epsilon \Rightarrow \ |x-a| \geq \delta\ (\epsilon > 0, \delta > 0)$ $\therefore P[|g(X_n) - g(a)| \geq \epsilon] \leq P[|X_n - a| \geq \delta]$
곱	$* X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$ ⑤ $\boldsymbol{X_nY_n \xrightarrow{P} XY}$	$* X_nY_n = \frac{1}{2}X_n^2 + \frac{1}{2}Y_n^2 - \frac{1}{2}(X_n - Y_n)^2 \xrightarrow{P} \frac{1}{2}X^2 + \frac{1}{2}Y^2 - \frac{1}{2}(X-Y)^2 = XY$

4. **일치성**: $\boldsymbol{T_n \xrightarrow{P} \theta}$ 면 ⇔ T_n은 θ의 **일치 추정량** $\quad * F(x;\theta)$에서 추출한 $iid\ \{X_1, \cdots, X_n\}$의 통계량 T_n

분산 추정량	① $\boldsymbol{S_n^2 \xrightarrow{P} \sigma^2}$ (일치 & 불편) ② $\boldsymbol{S_{mle}^2} = \frac{n-1}{n}S_n^2 \boldsymbol{\xrightarrow{P} \sigma^2}$ (일치 & MLE)
균등분포 모수 추정량	$X_1, \cdots, X_n \sim \text{unif}(0,\theta), \quad Y_n = \max\{X_1, \cdots, X_n\}$ $g_n(Y_n) = \frac{n!}{(n-1)!\,0!\,1!}F(Y_n)^{n-1}\left(1 - F(Y_n)\right)^0 f(y_n) = n\frac{t^{n-1}}{\theta^n} \quad (0 < t \leq \theta)$ $E(Y_n) = \int_0^\theta t\left(n\frac{t^{n-1}}{\theta^n}\right)dt = \frac{n}{n+1}\theta$ ∴ 최대값 Y_n은 θ의 일치 추정량
	$\bar{X}_n$은 $\theta/2$의 일치 추정량 ⇒ $2\bar{X}_n$은 θ의 일치 추정량

분포 수렴

1. **정의**: $\boldsymbol{X_n \xrightarrow{D} X}$ ⇔ $\forall x \in \{F_X$ 연속 점$\}$, $\boldsymbol{\lim_{n\to\infty} F_n(x) = F(x)}$, $\left(F: X\text{의 cdf},\ F_n: X_n\text{의 cdf}\right)$ "극한분포"

2. **t분포 ⇒ z분포** $(n \to \infty)$

   ① $f_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$, $\int_{-\infty}^{t} |f_n(y)|dy \leq \int_{-\infty}^{t} 10 f_1(y)dy = \frac{10}{\pi}\tan^{-1} t < \infty$ (르벡 DCT)

   ② $\lim_{n\to\infty} F_n(t) = \lim_{n\to\infty}\int_{-\infty}^{t} f_n(x)dx = \int_{-\infty}^{t}\lim_{n\to\infty} f_n(x)dx = \int_{-\infty}^{t}\lim_{n\to\infty}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}dx = \int_{-\infty}^{t}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)dx = \Phi(t)$

3. **정리**

① $\boldsymbol{X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X}$	https://freshrimpsushi.tistory.com/175?category=696570
② $\boldsymbol{X_n \xrightarrow{P} b \Leftrightarrow X_n \xrightarrow{D} b}$	if 분포수렴 ⇒ $\lim_{n\to\infty} P[|X_n - b| \leq \epsilon] = \lim_{n\to\infty} F_{X_n}(b+\epsilon) - F_{X_n}(b-\epsilon) = 1 - 0 = 0$
③ $\boldsymbol{X_n \xrightarrow{D} X}$ & $\boldsymbol{(A_n \xrightarrow{P} a,\ B_n \xrightarrow{P} b)}$ $\boldsymbol{\Rightarrow A_n + B_nX_n \xrightarrow{D} a + bX}$	**<Slutsky 정리>** e.g. $P_n - Q_n \xrightarrow{P} 0,\ Q_n \xrightarrow{D} X \Rightarrow P_n = (P_n - Q_n) + Q_n \xrightarrow{D} X$
* 받침 상 연속 $g(x)$ ④ $\boldsymbol{X_n \xrightarrow{D} X \Rightarrow g(X_n) \xrightarrow{D} g(X)}$	$\boldsymbol{Z_n \xrightarrow{D} Z \Rightarrow Z_n^2 \xrightarrow{D} \chi^2(1)}$
⑤ $\boldsymbol{X_n \xrightarrow{D} X \Leftrightarrow \lim_{n\to\infty} M_n(t) = M(t)}$	$\boldsymbol{Y_n \sim b(n,p) \Rightarrow \lim_{n\to\infty} M_n(t)} = \lim_{n\to\infty} E(e^{tY_n}) = \lim_{n\to\infty}[(1-p)+(pe^t)]^n = \boldsymbol{e^{\mu(e^t - 1)}}$ ∴ **이항분포 $\boldsymbol{b(n,p)} \xrightarrow{D}$ 푸아송분포** $(\boldsymbol{\mu = np})$

4. 일치성 / 극한분포 (“통계학적 수렴”)

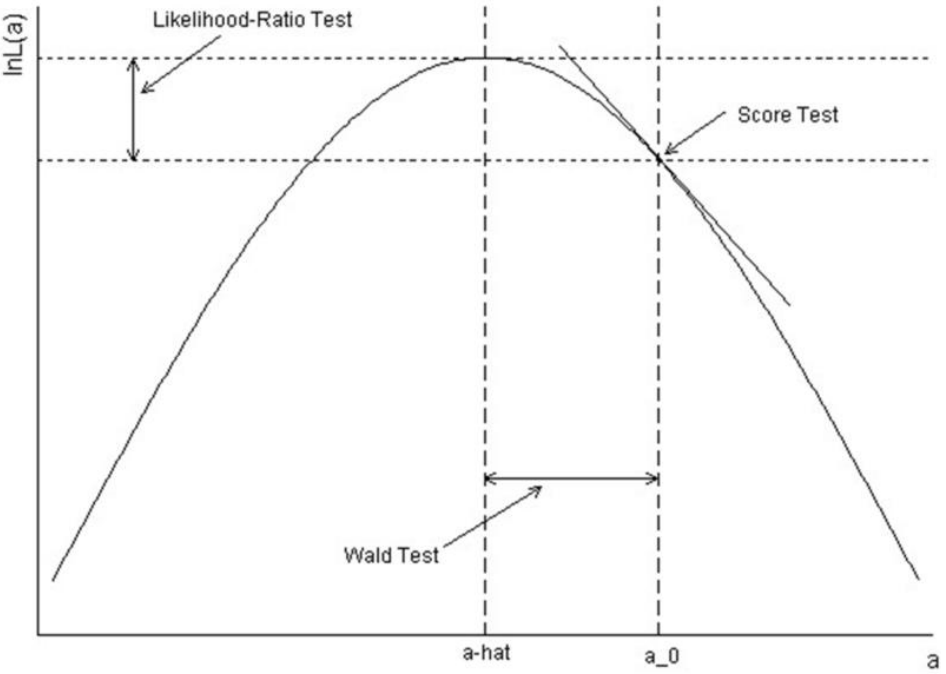
Δ-방법	$\sqrt{n}(X_n-\theta) \xrightarrow{D} N(0,\sigma^2)$ 이고, $g(x)$가 θ에서 미분 가능 & $g'(\theta)\neq 0$ 이면 $\sqrt{n}(g(X_n)-g(\theta)) \xrightarrow{D} N(0, g'(\theta)^2\sigma^2)$ **[Δ-method를 잘 이용하면 모수에 종속되지 않는 통계량 분산 만듦]** pf) 테일러 정리에 의해 $g(X_n)=g(\theta)+g'(\theta)(X_n-\theta)+o(\lvert X_n-\theta\rvert)$ 이므로 $\sqrt{n}\big(g(X_n)-g(\theta)\big)=\sqrt{n}g'(\theta)(X_n-\theta)+o(\sqrt{n}\lvert X_n-\theta\rvert) \xrightarrow{P} \sqrt{n}g'(\theta)(X_n-\theta) \xrightarrow{D} N(0,g'(\theta)^2\sigma^2)$ (중간에 little-o를 0으로 확률수렴 시키는 전개는 확률 유계인 Y_n에 대해 $o(Y_n)\xrightarrow{P}0$ 임을 이용)
중심극한정리 (CLT)	1. 중심극한정리: $Z_n=\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$ ← iid $X_i\sim$(평균: μ, 분산: σ^2) 2. 대표본 추론 통계량: $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \xrightarrow{D} N(0,1)$ ∵ $S\xrightarrow{P}\sigma \Leftrightarrow \dfrac{S}{\sigma}\xrightarrow{P}1$, CLT & Slutsky에 의해 $\left(\dfrac{\sigma}{S}\right)\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$
	2. 증명: MGF 이용 (특성함수 $\varphi(t)=E(e^{itx})$ 이용해야 더 정확함) $m(t):=E\left[e^{t(X-\mu)}\right]=e^{-\mu t}M(t) \Rightarrow m(0)=1,\ m'(0)=E(X-\mu)=0,\ m''(0)=E[(X-\mu)^2]+m'(0)^2=\sigma^2$ 테일러 정리에 의해 $m(t)=m(0)+m'(0)t+\frac{1}{2}m''(\xi)t^2=1+\frac{1}{2}m''(\xi)t^2=1+\frac{1}{2}\sigma^2t^2+\frac{1}{2}(m''(\xi)-\sigma^2)t^2,\quad \xi\in[-t,t]$ $M(t;n):=E(e^{tZ_n})=E\left(\exp\left(t\dfrac{(1/n)\sum X_i-\mu}{\sigma/\sqrt{n}}\right)\right)=E\left(\exp\left(t\dfrac{\sum_{i=1}^n(X_i-\mu)}{\sigma\sqrt{n}}\right)\right)=\prod_{i=1}^n E\left(\exp\left(t\dfrac{X_i-\mu}{\sigma\sqrt{n}}\right)\right)$ $=\left[E\left(\exp\left(\dfrac{t(X-\mu)}{\sigma\sqrt{n}}\right)\right)\right]^n=\left[m\left(\dfrac{t}{\sigma\sqrt{n}}\right)\right]^n,\quad -h<\dfrac{t}{\sigma\sqrt{n}}<h$ $M(t;n)=\left[m\left(\dfrac{t}{\sigma\sqrt{n}}\right)\right]^n=\left\{1+\dfrac{t^2}{2n}+\dfrac{[m''(\xi)-\sigma^2]t^2}{2n\sigma^2}\right\}^n,\quad \xi\in\left[-\dfrac{t}{\sigma\sqrt{n}},\dfrac{t}{\sigma\sqrt{n}}\right]$ $\therefore \lim_{n\to\infty}M(t;n)=\lim_{n\to\infty}\left\{1+\dfrac{t^2}{2n}+\dfrac{[m''(\xi)-\sigma^2]t^2}{2n\sigma^2}\right\}^n=\lim_{n\to\infty}\left(1+\dfrac{t^2}{2n}\right)^n=\exp\left(\dfrac{1}{2}t^2\right)\quad ∵ \lim_{n\to\infty}[m''(\xi)-\sigma^2]=0\ (∵ \xi\to 0)$ Z_n의 mgf $M(t;n)$의 $n\to\infty$ 극한값은 $N(0,1)$의 mgf $\exp\left(\frac{1}{2}t^2\right) \Rightarrow \therefore Z_n \xrightarrow{D} N(0,1)$

다변량 분포 확장	**수렴성 다변량 확장**	1) 확률수렴: $\{\mathbf{X_n}\}\in\mathbb{R}^p$일 때, 벡터의 각 성분이 수렴하는 경우가 전체 벡터의 수렴과 동치이다. 즉, $\mathbf{X_n}\xrightarrow{P}\mathbf{X} \Leftrightarrow X_{nj}\xrightarrow{P}X_j$ (모든 $j=1,\cdots,p$에서 성립) 2) 분포수렴: $\mathbf{X_n}\xrightarrow{D}\mathbf{X} \Leftrightarrow \forall\mathbf{x}\in\{F(\mathbf{x})$ 연속 점$\}$, $\lim_{n\to\infty}F_n(\mathbf{x})=F(\mathbf{x})$, $\big(F$: $\mathbf{X}$의 cdf, F_n: $\mathbf{X_n}$의 cdf$\big)$ ① $\mathbf{X_n}\xrightarrow{D}\mathbf{X}\Rightarrow g(\mathbf{X_n})\xrightarrow{D}g(\mathbf{X})$ (corollary: $g(\mathbf{x})=x_j$로 두면 분포수렴이 **주변 (marginal)** 수렴 수반 ② $\mathbf{X_n}\xrightarrow{D}\mathbf{X} \Leftrightarrow \lim_{n\to\infty}M_n(\mathbf{t})=M(\mathbf{t})$
	다변량 표본	$\{\mathbf{X_n}\}\in\mathbb{R}^p$ 인 평균 $\boldsymbol{\mu}$, 공분산행렬 $\boldsymbol{\Sigma}$ 인 iid 확률벡터열 ① 표본평균벡터: $\bar{\mathbf{X}}_n=\dfrac{1}{n}\sum_{i=1}^n\mathbf{X}_i=(\bar{X}_1,\cdots,\bar{X}_p)^T$ ② 표본공분산행렬: $S_j^2=\dfrac{1}{n-1}\sum_{i=1}^n(X_{ij}-\bar{X}_j)^2$, $S_{jk}=\dfrac{1}{n-1}\sum_{i=1}^n(X_{ij}-\bar{X}_j)(X_{ik}-\bar{X}_k)$; p × p 행렬 ∴ $\bar{\mathbf{X}}_n\xrightarrow{P}\boldsymbol{\mu}$, $\mathbf{S}_n\xrightarrow{P}\boldsymbol{\Sigma}$ (4차 적률 유한할 때 대수 약법칙)
	CLT	$\mathbf{Y_n}=\sqrt{n}(\bar{\mathbf{X}}_n-\boldsymbol{\mu})\xrightarrow{D}N_p(\mathbf{0},\boldsymbol{\Sigma}) \Leftrightarrow$ 근사적으로 $\bar{\mathbf{X}}_n\sim N_p\left(\boldsymbol{\mu},\frac{1}{n}\boldsymbol{\Sigma}\right)$
	Δ방법	$\sqrt{n}(\mathbf{X_n}-\boldsymbol{\mu_0})\xrightarrow{D}N_p(\mathbf{0},\boldsymbol{\Sigma})$ ($\mathbf{g}$는 $\mathbb{R}^p\to\mathbb{R}^k$ 로의 변환 $(k\le p)$; 미분행렬 $\mathbf{B}=\left[\dfrac{\partial g_i}{\partial x_j}\right]$이 연속, $\mathbf{B}\neq\mathbf{0}$ in $\boldsymbol{\mu_0}$ 근방 $\sqrt{n}(\mathbf{g}(\mathbf{X_n})-\mathbf{g}(\boldsymbol{\mu_0}))\xrightarrow{D}N_p(\mathbf{0},\mathbf{B_0}\boldsymbol{\Sigma}\mathbf{B_0}^T)$ $\mathbf{B_0}=\mathbf{B}(\boldsymbol{\mu_0})$

5. 최대가능도방법 (Maximum Likelihood Methods)

MLE **(R0)~(R2)**	**MLE 핵심**	(R0), (R1) 하에서 $\lim_{n\to\infty} P_{\theta_0}[L(\theta_0,\mathbf{X}) > L(\theta,\mathbf{X})] = 1 \quad (\forall\,\theta \neq \theta_0)$ $pf)\ \frac{1}{n}\sum_{i=1}^{n}\ln\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right] \xrightarrow{P} E_{\theta_0}\left(\ln\left[\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right]\right) < \ln E_{\theta_0}\left[\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right]$ by 대수의 법칙, 젠센 부등식 $E_{\theta_0}\left[\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right] = \int \frac{f(x;\theta)}{f(x;\theta_0)} f(x;\theta_0)dx = 1$ ($R1$ 공통 받침 하에서) $\therefore\ \frac{1}{n}\sum_{i=1}^{n}\ln\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right] < 0 \Leftrightarrow\ L(\theta_0,\mathbf{X}) > L(\theta,\mathbf{X})$ ∴ 근사적으로 **참값 $\boldsymbol{\theta_0}$에서 우도함수 $\boldsymbol{L(\theta,\mathbf{X})}$가 최대**가 된다. ($\hat\theta = \mathrm{Argmax}[L(\theta)] \xrightarrow{P} \theta_0$)
	불변성	$\boldsymbol{\eta = g(\theta) \Leftrightarrow \hat\eta = g(\hat\theta)}$ $pf)$ ① $g \in$ 1대1 함수: $\max L(\theta) = \max L\left(g^{-1}(\eta)\right)$ 이므로 $\hat\theta = g^{-1}(\hat\eta)$에서 우도 최대화 ② $g \notin$ 1대1 함수: $g^{-1}(\eta) := \{\theta: g(\theta) = \eta\}$ 새로 정의 → $\hat\theta \in g^{-1}(\hat\eta)$ 에서 우도최대화
	추정 방정식	*추정방정식 (estimating equation; EE): $\partial l(\theta)/\partial\theta = 0$ (R0)~(R2) 하에서 $\partial l(\theta)/\partial\theta = 0$ 는 $\hat\theta \xrightarrow{P} \theta_0$ 인 $\hat\theta$를 가짐 (Corollary: EE가 유일해를 가지면 그 해는 $\hat\theta \xrightarrow{P} \theta_0$)
Cramér Rao Bound **(R0)~(R4)**	**스코어 함수 & 피셔정보**	① **Score 함수** $s(\theta) = \frac{\partial \ln f}{\partial\theta}$ ② **Fisher information** $I(\theta) = \mathrm{Var}\left(\frac{\partial \ln f}{\partial\theta}\right) = \mathrm{E}\left[\left(\frac{\partial \ln f}{\partial\theta}\right)^2\right] = -\mathrm{E}\left[\frac{\partial^2 \ln f}{\partial\theta^2}\right]$
		$1 = \int_{-\infty}^{\infty} f dx$ → 양변 θ로 i) 한번 미분 ii) 두번 미분 하면 i) $0 = \int_{-\infty}^{\infty}(\partial f/\partial\theta)dx = \int_{-\infty}^{\infty}\frac{(\partial f/\partial\theta)}{f} f dx = \int_{-\infty}^{\infty}\left(\frac{\partial \ln f}{\partial\theta}\right) f dx \qquad \therefore\ \boldsymbol{E\left(\frac{\partial \ln f}{\partial\theta}\right) = 0}$ ii) $0 = \int_{-\infty}^{\infty}\frac{\partial^2 \ln f}{\partial\theta^2} f dx + \int_{-\infty}^{\infty}\left(\frac{\partial \ln f}{\partial\theta}\right)\left(\frac{\partial \ln f}{\partial\theta}\right) f dx \quad \therefore\ \boldsymbol{\mathrm{E}\left[\frac{\partial^2 \ln f}{\partial\theta^2}\right] + \mathrm{E}\left[\left(\frac{\partial \ln f}{\partial\theta}\right)^2\right] = 0}$
		iid $[X_1,\cdots,X_n]$ 에 대해서 ① **Score 함수** $s_n(\theta) = \frac{\partial l}{\partial\theta} = \frac{\partial \ln L}{\partial\theta} = \sum_{i=1}^{n}\frac{\partial \ln f(X_i;\theta)}{\partial\theta}$ ② **Fisher 정보** $I_n(\theta) = \mathrm{Var}\left(\frac{\partial l}{\partial\theta}\right) = \mathrm{Var}\left(\frac{\partial \ln L}{\partial\theta}\right) = n\,I(\theta)$
	Cramér-Rao Bound (CRB)	① $\boldsymbol{\mathrm{Var}(T) \geq \frac{[\partial E(T)/\partial\theta]^2}{nI(\theta)}}$ for 임의의 통계량 $T = g(X_1,\cdots,X_n)$ ② $\boldsymbol{\mathrm{Var}(T) \geq \frac{1}{nI(\theta)}}$ for 불편추정량 T ($\because\ E(T) = \theta$)
		$pf)\ E(T) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}[T]f(x_1;\theta)\cdots f(x_n;\theta)dx_1\cdots dx_n$ $\Leftrightarrow\ \partial E(T)/\partial\theta = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}[T]\left[\sum_{i=1}^{n}\frac{\partial \ln f(x_i;\theta)}{\partial\theta}\right] f(x_1;\theta)\cdots f(x_n;\theta)\,dx_1\cdots dx_n \quad ,Z := \left[\sum_{i=1}^{n}\frac{\partial \ln f(x_i;\theta)}{\partial\theta}\right]$ $\Leftrightarrow\ \partial E(T)/\partial\theta = E(TZ) = E(T)E(Z) + \rho\,\sigma_T\sigma_Z = \rho\sqrt{\mathrm{Var}(T)}\sqrt{nI(\theta)} \qquad \therefore\ \rho^2 \leq 1 \Leftrightarrow \mathrm{Var}(T) \geq \frac{[\partial E(T)/\partial\theta]^2}{nI(\theta)}$
	효율성	*효율성: 통계량 T의 효율성은 CRB(T)/Var(T) $*$ ARE (근사 상대효율성) $= \mathrm{e(T,W)} = \lim_{n\to\infty}\frac{\mathrm{Var(W)}}{\mathrm{Var(T)}}$ (if $\mathrm{T} \xrightarrow{P} \theta_0, \mathrm{W} \xrightarrow{P} \theta_0$일 때)
	MLE 정규근사 (R0)~(R5)	① **정규 근사**: $\boldsymbol{\sqrt{n}(\hat\theta - \theta_0) \xrightarrow{D} N\left(0, \frac{1}{I(\theta_0)}\right)}$ for 유한 피셔정보 $\mathrm{I}(\theta_0)$ $\quad *\,pf)\ l'(\hat\theta)$를 θ_0 테일러 전개 ... ➔ **<u>MLE의 근사 정규 신뢰 구간</u> 구할 수 있음.** $\therefore\ \mathrm{Var}(\hat{\boldsymbol\theta}) \xrightarrow{P} \boldsymbol{\frac{1}{nI(\theta_0)}}$ (mle는 근사적으로 효율적 or mle의 분산은 CRB에 근사) ② **Δ방법**: $\boldsymbol{\sqrt{n}\left(g(\hat\theta) - g(\theta_0)\right) \xrightarrow{D} N\left(0, \frac{g'(\theta_0)^2}{I(\theta_0)}\right)}$ ($g(x)$가 θ에서 미분 가능 & $g'(\theta) \neq 0$ 이면) ③ 정규 근사: $\hat\theta - \theta_0 = \frac{1}{nI(\theta_0)}\sum_{i=1}^{n}\frac{\partial \ln f(X_i;\theta_0)}{\partial\theta} + \frac{R_n}{\sqrt{n}} = -\frac{l'(\theta_0)}{l''(\theta_0)} + \frac{R_n}{\sqrt{n}} \quad \left(R_n \xrightarrow{P} 0\right)$
	MLE Newton's	$\hat\theta_1 = \hat\theta_0 - \frac{l'(\hat\theta_0)}{l''(\hat\theta_0)}$ 과정 반복 $\qquad *\ \hat\theta_0$이 일치 추정량이면 $\hat\theta_1$은 mle 임 $\left(\xrightarrow{D} \boldsymbol{N\left(0, \frac{1}{I(\theta_0)}\right)}\right)$

5. 최대가능도방법 (Maximum Likelihood Methods)

<table>
<tr><td rowspan="6">최대
가능도
검정
(ML tests)</td><td>전개</td><td colspan="2">우도비 (LR): $\boldsymbol{\Lambda = \frac{L(\theta_0)}{L(\hat\theta)}}$ ($\boldsymbol{\Lambda \le c}$ 에서 기각)
$-\frac{1}{n} l''(\theta_0) \xrightarrow{P} I(\theta_0)$, $\frac{l'(\theta_0)}{\sqrt{n}} = \sqrt{n}(\hat\theta - \theta_0) I(\theta_0) + R_n$ 이므로
$l(\hat\theta) = l(\theta_0) + (\hat\theta - \theta_0) l'(\theta_0) + \frac{1}{2}(\hat\theta - \theta_0)\, l''(\theta_n^*)$ 의 미분계수 항들에 대입해주면
$-2\ln\Lambda = 2[l(\hat\theta) - l(\theta_0)] = \left[\sqrt{nI(\theta_0)}(\hat\theta - \theta_0)\right]^2 + R_n^*$ $\left(R_n^* \xrightarrow{P} 0\right)$
∴ $\boldsymbol{-2\ln\Lambda \xrightarrow{D} \chi^2(1)}$ ⇐ $\sqrt{nI(\theta_0)}(\hat\theta - \theta_0) \xrightarrow{D} N(0,1)$</td></tr>
<tr><td>우도비 검정</td><td>$\chi_L^2 = \boldsymbol{-2\ln\Lambda}$</td><td rowspan="3">$\chi^2 \ge \chi_\alpha^2(1)$ 에서 단측 검정 기각역
$(H_0: \theta = \theta_0, H_1: \theta \ne \theta_0)$</td></tr>
<tr><td>Wald 검정</td><td>$\chi_W^2 = \left[\sqrt{nI(\hat\theta)}(\hat\theta - \theta_0)\right]^2$</td></tr>
<tr><td>Score 검정</td><td>$\chi_R^2 = \left(\frac{l'(\theta_0)}{\sqrt{nI(\theta_0)}}\right)^2$</td></tr>
<tr><td colspan="3">

have the following relationship Wald ≥ LR ≥ score (Johnston and DiNardo 1997 p. 150). That is, the Wald test statistic will always be greater than the LR test statistic, which will, in turn, always be greater than the test statistic from the score test. When computing power was much more limited, and many models took a long time to run, being able to approximate the LR test using a single model was a fairly major advantage. Today, for most of the models researchers are likely to want to compare, computational time is not an issue, and we generally recommend running the likelihood ratio test in most situations. This is not to say that one should never use the Wald or score tests. For example, the Wald test is commonly used to perform multiple degree of freedom tests on sets of dummy variables used to model categorical predictor variables in regression (for more information see our webbooks on Regression with Stata, SPSS, and SAS, specifically Chapter 3 – Regression with Categorical Predictors.) The advantage of the score test is that it can be used to search for omitted variables when the number of candidate variables is large.</td></tr>
<tr><td colspan="3"></td></tr>
<tr><td>정칙
조건

Regularity conditions</td><td colspan="3">(R0): pdf $f(x;\theta)$는 서로 distinct 하다. i.e. $\theta_1 \ne \theta_2 \Rightarrow f(x_i;\theta_1) \ne f(x_i;\theta_2)$
(R1): pdf $f(x;\theta)$는 모든 θ에 대해 공통된 support를 갖는다. (θ에 의존적이지 않다.)
(R2): θ_0 (참값) $\in \Omega$
(R3): pdf $f(x;\theta)$는 θ로 두 번 미분 가능
(R4): $\int f(x;\theta)dx$는 θ로 두 번 미분 가능
(R5): pdf $f(x;\theta)$는 θ로 세 번 미분 가능, 모든 θ에 대해 $\lvert\partial^3 \ln f/\partial\theta^3\rvert \le M(x)$ $\left(E_{\theta_0}[M(X)] < \infty\right)$ $in\ \theta_0$근방 $\forall x$</td></tr>
</table>

5. 최대가능도방법 (Maximum Likelihood Methods)

다중 모수

*정칙조건 ~(R9) 까지 추가됨. (기존 정칙의 다변량 확장)

맨위 "MLE 핵심" 정리는 벡터 $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_p]^T \in \mathbb{R}^p$에 대해서도 똑같이 성립함. ⇔ $\nabla l(\boldsymbol{\theta}) = \mathbf{0}$의 해 구하기

다중 모수 추정	**피셔정보량**	$\nabla \ln f(X;\boldsymbol{\theta}) = \left(\frac{\partial \ln f(X;\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \ln f(X;\boldsymbol{\theta})}{\partial \theta_p}]\right)^T$ 피셔 정보량: $\mathbf{I}(\boldsymbol{\theta}) = \mathrm{Cov}(\nabla \ln f(X;\boldsymbol{\theta})) = -E\left[\frac{\partial^2}{\partial \theta_j \partial \theta_k} \ln f\right]_{jk} = E\left[\left(\frac{\partial \ln f}{\partial \theta_j}\right)\left(\frac{\partial \ln f}{\partial \theta_k}\right)\right]_{jk}$
	피셔정보량 (표본 n 확장)	$\nabla l = \nabla \ln L = \sum_{i=1}^{n} \nabla \ln f$ 피셔 정보량: $\mathbf{I_n}(\boldsymbol{\theta}) = \mathrm{Cov}(\nabla l) = \mathrm{Cov}(\nabla \ln L) = n\mathbf{I}(\boldsymbol{\theta})$
	CRB	$Var(T_j) \geq \frac{1}{n}[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{jj}$ (T_j가 θ_j의 불편 추정량)
	정규근사	$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0}) \xrightarrow{D} N_p(\mathbf{0}, \mathbf{I}^{-1}(\boldsymbol{\theta_0}))$ ⇒ $\sqrt{n}(\hat{\theta}_j - \theta_j) \xrightarrow{D} N(0, [\mathbf{I}^{-1}(\boldsymbol{\theta_0})]_{jj})$
	Δ방법	$\sqrt{n}(\mathbf{g}(\hat{\boldsymbol{\theta}}) - \mathbf{g}(\boldsymbol{\theta_0})) \xrightarrow{D} N_p(\mathbf{0}, \mathbf{B}[\mathbf{I}^{-1}(\boldsymbol{\theta_0})]\mathbf{B}^T)$ ($\mathbf{g}$는 $\mathbb{R}^p \to \mathbb{R}^k$ 로의 변환 ($k \leq p$); 미분행렬 $\mathbf{B} = \left[\frac{\partial g_i}{\partial \theta_j}\right]$이 연속, $\mathbf{B} \neq \mathbf{0}$ in $\boldsymbol{\theta_0}$ 근방)

다중 모수 검정

기본

$\boldsymbol{H_0}: \boldsymbol{\theta} \in \boldsymbol{\omega}$, $\boldsymbol{H_1}: \boldsymbol{\theta} \in (\boldsymbol{\omega}^c \cap \boldsymbol{\Omega})$ *예시: 전체 자유모수 k-1 & H_0 자유모수 0 ➔ q=k-1

(Ω: p차원 전체 모수 공간; ω: p-q차원 귀무가설의 모수공간 (자유도))

우도비 (LR): $\boldsymbol{\Lambda} = \frac{\boldsymbol{L(\hat{\omega})}}{\boldsymbol{L(\hat{\Omega})}} = \frac{\max_{\boldsymbol{\theta} \in \boldsymbol{\omega}} \boldsymbol{L(\theta)}}{\max_{\boldsymbol{\theta} \in \boldsymbol{\Omega}} \boldsymbol{L(\theta)}}$

$\chi_L^2 = -2\ln \boldsymbol{\Lambda} \xrightarrow{D} \chi^2(q)$ (Wald, Score 검정통계량도 가능)

예시

정규 μ

$\boldsymbol{H_0}: \boldsymbol{\mu} = \boldsymbol{\mu_0}, \boldsymbol{H_1}: \boldsymbol{\mu} \neq \boldsymbol{\mu_0}$ $\{\boldsymbol{X_n}\} \overset{\text{iid}}{\sim} \boldsymbol{N(\mu, \sigma^2)}$

$$L(\hat{\Omega}) = \frac{1}{(2\pi\hat{\sigma}^2)^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\frac{1}{n}(\sum_{i=1}^{n}(X_i - \bar{X})^2)}\right\} = \frac{\mathbf{1}}{(\mathbf{2}\boldsymbol{\pi}\hat{\boldsymbol{\sigma}})^{\boldsymbol{n/2}}} \mathbf{exp}\left(-\frac{\boldsymbol{n}}{\mathbf{2}}\right)$$

$$L(\hat{\omega}) = \frac{\mathbf{1}}{(\mathbf{2}\boldsymbol{\pi}\hat{\sigma}_0^2)^{\boldsymbol{n/2}}} \mathbf{exp}\left(-\frac{\boldsymbol{n}}{\mathbf{2}}\right) \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2, \quad \hat{\sigma}_0{}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$$

$$\left(\frac{1}{\Lambda}\right)^{\frac{2}{n}} = \left(\frac{L(\hat{\Omega})}{L(\hat{\omega})}\right)^{\frac{2}{n}} = \frac{\sum_{i=1}^{n}(X_i - \mu_0)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = 1 + \frac{n(\bar{X} - \mu_0)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = 1 + \frac{1}{n-1}\left\{\frac{(\bar{X} - \mu_0)}{S/\sqrt{n}}\right\}^2$$

$\left(\frac{1}{\Lambda}\right)^{\frac{2}{n}} \geq c' \Leftrightarrow |T| \geq c^* = \sqrt{(c'-1)(n-1)}$ ∴ 양측 t검정과 동치

다항 p

$\boldsymbol{H_0}: \boldsymbol{p_1} = \boldsymbol{p_2}$, $\boldsymbol{H_1}: \boldsymbol{p_1} \neq \boldsymbol{p_2}$ (유력후보1 vs 유력후보2 vs 나머지 군소후보)

3항 베르누이 $(X_{i1}, X_{i2}) \sim \boldsymbol{p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{1 - x_1 - x_2}}$ $(X_{i1}, X_{i2}) \in \{(0,0), (0,1), (1,0)\}$

$\hat{p}_j = \frac{\sum_{i=1}^{n} X_{ij}}{n}$ for $j = 1,2$ (표본: $\{(X_{n1}, X_{n2})\}$)

LR	$\frac{1}{\Lambda} = \left(\frac{2\hat{p}_1}{\hat{p}_1 + \hat{p}_2}\right)^{n\hat{p}_1} \left(\frac{2\hat{p}_2}{\hat{p}_1 + \hat{p}_2}\right)^{n\hat{p}_2}$, $-2\ln\Lambda > \chi_\alpha^2(1)$ 에서 기각
Wald	$\begin{bmatrix}\hat{p}_1 \\ \hat{p}_2\end{bmatrix} \overset{a}{\sim} N_2\left(\begin{bmatrix}p_1 \\ p_2\end{bmatrix}, \frac{1}{n}\begin{bmatrix}p_1(1-p_1) & -p_1p_2 \\ -p_1p_2 & p_2(1-p_2)\end{bmatrix}\right)$ $W = \hat{p}_1 - \hat{p}_2 = g\left(\begin{bmatrix}\hat{p}_1 \\ \hat{p}_2\end{bmatrix}\right)$, Δ방법에서 $\mathbf{B} = \left[\frac{\partial g_i}{\partial p_j}\right] = [1, -1]$ $Var(W) = \frac{1}{n}\mathbf{B}\mathbf{I^{-1}}\mathbf{B}^T = \frac{p_1 + p_2 - (p_1 - p_2)^2}{\mathrm{n}}$ $\therefore W \overset{a}{\sim} N\left(p_1 - p_2, \frac{p_1 + p_2 - (p_1 - p_2)^2}{\mathrm{n}}\right)$ ⇒ 근사 Z 검정 or 카이제곱

2표본 이항 p

$\boldsymbol{H_0}: \boldsymbol{p_1} = \boldsymbol{p_2}$, $\boldsymbol{H_1}: \boldsymbol{p_1} \neq \boldsymbol{p_2}$, $\{\boldsymbol{X_{n1}}\} \overset{\text{iid}}{\sim} \boldsymbol{B(1, p_1)}$, $\{\boldsymbol{Y_{n2}}\} \overset{\text{iid}}{\sim} \boldsymbol{B(1, p_2)}$

Wald	$\hat{p}_1 \overset{a}{\sim} N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right)$, $\hat{p}_2 \overset{a}{\sim} N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$, $\mathrm{Cov}(\hat{p}_1, \hat{p}_2) = 0$ $\hat{p}_1 - \hat{p}_2 \overset{a}{\sim} N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$ & $Slutsky$ $(\hat{p}_1 \xrightarrow{P} p_1, \hat{p}_2 \xrightarrow{P} p_2)$ ⇒ 근사 Z 검정 or 카이제곱

6. 충분성 (Sufficiency) – 통계량의 성질

통계량 Review	**통계량**	① 점추정: $\boldsymbol{\theta} \in \Omega$**에 대한 추정량** $\hat{\boldsymbol{\theta}}$ *통계량 (Statistic): $T = T(X_1, \cdots, X_n)$ (표본에 대한 함수) ② 95% CI: $0.95 = P_\theta[\theta \in (\hat{\theta}_L, \hat{\theta}_U)]$ *$\theta \in (\hat{\theta}_L, \hat{\theta}_U)$인 베르누이 사건 ~ $B(1, 0.95)$
	성질	1) 일치추정량: $\boldsymbol{T_n \xrightarrow{P} \theta}$ 면 ⇔ T_n은 θ의 **일치 추정량**
		2) 불편추정량: $E(T) = \theta$ ⇔ T는 θ의 불편 추정량 (bias = 0) ① **MVUE**: 분산 최소인 불편추정량 (UE) ➔ 유일 ② **CRB**: $\mathbf{Var}(\boldsymbol{T}) \geq \mathbf{1}/\{\boldsymbol{nI(\theta)}\}$
		3) MLE: $\hat{\theta} = \mathrm{Argmax}[L(\theta)] = \mathrm{Argmax}[\prod_{i=1}^n f(x_i, \theta)]$ ① MLE는 근사적으로 효율적 ② $\hat{\boldsymbol{\theta}} \overset{a}{\sim} \boldsymbol{N}\left(\boldsymbol{\theta_0}, \frac{\mathbf{1}}{\boldsymbol{nI(\theta_0)}}\right)$ ⇒ Z or χ^2화 하면 Wald statistic
		4) $ARE(T_1, T_2) = \dfrac{Var(T_2)}{Var(T_1)}$
	Bias MSE	1) **bias** $(\hat{\boldsymbol{\theta}}) = \boldsymbol{E}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = E(\hat{\theta}) - \theta$ $*$ $\mathrm{bias}(g(\hat{\theta})) = E\left(g(\hat{\theta}) - g(\theta)\right)$ $*$ $bias \xrightarrow{P} 0$, if θ가 일치 추정량 2) Mean square error (MSE): $\boldsymbol{mse}(\hat{\boldsymbol{\theta}}) = \boldsymbol{E}\left\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^2\right\} = Var(\hat{\theta}) + \{\mathrm{bias}(\hat{\theta})\}^2$ 3) Mean absolute error (MAE): $\boldsymbol{mse}(\hat{\boldsymbol{\theta}}) = \boldsymbol{E}\{\lvert\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\rvert\}$
	적률 추정법 (MoM)	r차 표본적률 $\xrightarrow{P}$ r차 모적률 (➔ 연립하여 모수 추정량 구함; 일반적으로 비선호)
		ex) $\{X_i\} \overset{\text{iid}}{\sim} \mathrm{Gamma}(k, \theta)$ $m_1 = \dfrac{\sum_{i=1}^n X_i}{n} = \hat{k}\hat{\theta}$, $m_2 = \dfrac{\sum_{i=1}^n X_i^2}{n} = \hat{k}(\hat{\theta})^2 + (\hat{k}\hat{\theta})^2$ $\hat{\theta} = \dfrac{\sum_{i=1}^n (X_i - \bar{X})^2}{n\bar{X}} = \left(\dfrac{n-1}{n}\right)\dfrac{S^2}{\bar{X}} = \dfrac{S_{mle}^2}{\bar{X}}$, $\hat{k} = \dfrac{n(\bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} = \left(\dfrac{n}{n-1}\right)\dfrac{\bar{X}}{S^2} = \dfrac{\bar{X}^2}{S_{mle}^2}$
충분성	**정의**	$Y = u(X_1, \cdots, X_n)$에 대해 $\boldsymbol{X}\vert\boldsymbol{y}$**가 θ와 무관**함 ⇔ Y가 θ에 대한 모든 정보 다 포함 (e.g. $Y = \Sigma_{i=1}^n X_i$) $\dfrac{\prod_{i=1}^n f(x_i, \theta)}{f_Y(y; \theta)} = H(x_1, \cdots, x_n)$ (f_Y: Y의 pdf)
	Neyman-Fisher	[Y가 θ의 SS] ⇔ $\prod_{i=1}^n f(x_i, \theta) = f_Y(y; \theta)\, H(x_1, \cdots, x_n) = g(y; \theta)\, h(x_1, \cdots, x_n)$ (임의의 g, h로 **인수분해**)
	Rao-Blackwell	θ의 충분통계량 Y_1, 불편추정량 Y_2에 대해, 새로운 불편추정량 $\boldsymbol{\varphi(y_1) = E(Y_2 \vert y_1)}$ $E(\varphi(Y_1)) = E[E(Y_2\vert Y_1)] = E(Y_2) = \theta$ 2) $\mathrm{Var}(\varphi(Y_1)) = \mathrm{Var}(E(Y_2\vert Y_1)) \leq \mathrm{Var}(Y_2)$ ∴ **New** $\boldsymbol{UE}$ $\varphi(y_1) = E(Y_2\vert y_1)$는 **Old** $\boldsymbol{UE}$ Y_2 보다 분산이 작다. *실전: $E(\varphi(Y_1)) = \theta$인 $\varphi(y_1)$찾기
	Lehmann-Scheffe	① 통계량 Y는 **complete (완비)** if 모든 θ에서 $E(h(Y)) = 0$ ⇒ $h(t) = 0$ 만 가능함 ② **레만-셰페: CSS**인 $\boldsymbol{Y_1}$으로 Rao-Blackwellization ➔ $\boldsymbol{\varphi(y_1) = E(Y_2\vert y_1)}$는 유일한 $\boldsymbol{MVUE}$ of $\boldsymbol{\theta}$ pf) CSS인 Y_1에 대해 불편추정량 $\varphi(Y_1), \psi(Y_1)$존재 ⇒ $E(\varphi(Y_1) - \psi(Y_1)) = \theta - \theta = 0$ 완비족 $\{f_{Y_1}(y; \theta): \theta \in \Omega\}$에 대해 위 등식은 $\varphi(Y_1) = \psi(Y_1)$에서만 성립 (더 이상 분산 못 줄임)
	지수족 Exponential Family	$f(x; \theta) = \exp[\eta(\theta)T(x) + H(x) - A(\eta(\theta))]$ $(x \in S)$ ($\eta = \eta(\theta)$는 자연 모수) *정칙: 1) $\boldsymbol{S}$**가** $\boldsymbol{\theta}$**에 종속 X**, 2) $\boldsymbol{\eta(\theta)}$**연속**, 3) (연속이면) $\boldsymbol{H(x)}$**연속** in $\{K'(x) \neq 0\}$ ① 지수족: 이산 (포아송, 이항, 기하, 음이항, 다항 등) / 연속 (감마, 베타, 정규 등) ② $Y = \sum_{i=1}^n T(x)$ 는 θ의 **CSS** ③ $E(T(X)) = A'(\eta)$, $\mathrm{Var}(T(X)) = A''(\eta)$
	결합 충분통계량 (다중모수)	$\mathbf{Y} = (Y_1, \cdots, Y_m)^T \in \mathbb{R}^m$ & $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p) \in \mathbb{R}^p$에 대해 (일반적으로 m = p) $\prod_{i=1}^n f(x_i, \boldsymbol{\theta}) = f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\theta})H(x_1, \cdots, x_n) = g(\mathbf{y}; \boldsymbol{\theta})\, h(x_1, \cdots, x_n)$ *순서통계량 $\mathbf{Y} = (Y_1, \cdots, Y_n)^T$; $Y_1 < \cdots < Y_n$ → 모든 연속분포의 결합충분통계량
	보조통계량 (Ancillary)	$A = a(X_1, \cdots, X_n)$가 θ와 **무관** ex) 정규분포 iid의 $\boldsymbol{S^2}$: $\boldsymbol{\mu}$에 대해 ancillary 1) **Basu** 정리:{$\boldsymbol{Y}$가 θ 의 **CSS**} & {$\boldsymbol{Z}$가 θ 의 **ancillary**} ⇔ {$\boldsymbol{Y}$와 $\boldsymbol{Z}$는 **독립**} ex) $\bar{\boldsymbol{X}} \perp \boldsymbol{S^2}$, $\{X_i\} \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ 2) ①**위치불변**:$Z = u(W_1 + \theta, \cdots, W_n + \theta) = u(W_1, \cdots, W_n)$ ex) S^2, $\max\{X_i\} - \min\{X_i\}$, $\frac{1}{n}\sum_1^n \lvert X_i - Q_2\rvert$ ②**척도불변**: $Z = u(\theta W_1, \cdots, \theta W_n) = u(W_1, \cdots, W_n)$ ex) $X_1/(X_1 + X_2)$, $X_1^2/\sum_1^n X_i^2$, $\min\{X_i\}$ / $\max\{X_i\}$ ③**위치척도불변**: $Z = u(\theta_1 W_1 + \theta_2, \cdots, \theta_1 W_n + \theta_2) = u(W_1, \cdots, W_n)$ ex) $(X_i - \bar{X})/S^2$
	MLE	$L(\theta; x_1, \cdots, x_n) = \prod_{i=1}^n f(x_i, \theta) = f_Y(y; \theta)\, H(x_1, \cdots, x_n)$ ➔ L과 f_y 동시에 극대화 by θ ① MLE $\hat{\theta}$이 유일 ⇔ $\hat{\theta}$는 충분통계량 Y의 함수 ∵ $\hat{\theta} = \mathrm{argmax}(L(\theta, \mathbf{x})) = \mathrm{argmax}(f_Y(y; \theta))$ ② MLE $\hat{\theta}$가 충분통계량 ⇔ $\hat{\theta}$는 최소 충분통계량 (MSS) *최소충분:reduced from 다른 충분통계량

6. 충분성 (Sufficiency) – 통계량의 성질

지수족 확장

지수족

$f(x;\theta) = \exp[\eta(\theta)T(x) + H(x) - A(\eta(\theta))]$ $(x \in S)$ $(\eta = \eta(\theta)$는 자연 모수)

*정칙: 1) **S가 θ에 종속 X**, 2) **$\eta(\theta)$연속**, 3) (연속이면) **$H(x)$연속** in $\{K'(x) \neq 0\}$

① 지수족: 이산 (포아송, 이항, 기하, 음이항, 다항 등) / 연속 (감마, 베타, 정규 등)

② $Y = \sum_{i=1}^{n} T(x)$ 는 θ의 **CSS** ③ $E\big(T(X)\big) = A'(\eta)$, $\mathrm{Var}\big(T(X)\big) = A''(\eta)$

다변량 확장

1변수 1모수	$f(x;\theta) = \exp[\eta(\theta)T(x) + H(x) - A(\eta)]$
1변수 다중모수	$f(x;\boldsymbol{\theta}) = \exp[\boldsymbol{\eta}(\boldsymbol{\theta}) \cdot \mathbf{T}(x) + H(x) - A(\boldsymbol{\eta})]$
다변량 다중모수	$f(\mathbf{x};\boldsymbol{\theta}) = \exp[\boldsymbol{\eta}(\boldsymbol{\theta}) \cdot \mathbf{T}(\mathbf{x}) + H(\mathbf{x}) - A(\boldsymbol{\eta})]$
기대값	$\nabla A(\boldsymbol{\eta}) = E[\mathbf{T}(\mathbf{x})]$, $\mathbf{H}[A(\boldsymbol{\eta})] = \mathrm{Cov}(\mathbf{T}(\mathbf{x}))$ $\nabla A(\boldsymbol{\eta}_{mle}) = \dfrac{1}{n}\sum_{i=1}^{n} \mathbf{T}(\mathbf{x}_i)$

예시

분포	모수 $\boldsymbol{\theta}$	자연모수 $\boldsymbol{\eta}$	역모수	$\mathbf{T}(\mathbf{x})$	$A(\boldsymbol{\eta})$
베르누이	p	$\ln\dfrac{p}{1-p}$	$\dfrac{1}{1+e^{-\eta}}$ * logistic function	x	$\ln(1+e^{\mu})$
이항					$n\ln(1+e^{\mu})$
푸아송	m	$\ln m$	e^{η}	x	e^{η}
음이항(r)	p	$\ln(1-p)$	$1-e^{\eta}$	x	$-r\ln(1-e^{\mu})$
다항(n)	$\begin{bmatrix} p_1 \\ \vdots \\ p_{k-1} \end{bmatrix}$	$\begin{bmatrix} \ln\dfrac{p_1}{p_k} \\ \vdots \\ \ln\dfrac{p_{k-1}}{p_k} \end{bmatrix}$	$\begin{bmatrix} \dfrac{\exp(\eta_1)}{1+\sum_{j=1}^{k-1}\exp(\eta_j)} \\ \vdots \\ \dfrac{\exp(\eta_{k-1})}{1+\sum_{j=1}^{k-1}\exp(\eta_j)} \end{bmatrix}$ * softmax function	$\begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix}$	$n\ln(1+\sum_{j=1}^{k-1}\exp(\eta_j))$
	$p_k = 1-\sum_{j=1}^{k-1}p_j$, $\eta_k = 0$, $\exp(\eta_k) = 1$				
감마	$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$	$\begin{bmatrix} \alpha-1 \\ -\dfrac{1}{\beta} \end{bmatrix}$	$\begin{bmatrix} \eta_1+1 \\ -\dfrac{1}{\eta_2} \end{bmatrix}$	$\begin{bmatrix} \ln x \\ x \end{bmatrix}$	$\ln\Gamma(\eta_1+1) - (\eta_1+1)\ln(-\eta_2)$
지수	β	$-\dfrac{1}{\beta}$	$-\dfrac{1}{\eta}$	x	$-\ln(-\eta)$
카이제곱	ν	$\dfrac{\nu}{2}-1$	$2(\eta+1)$	$\ln x$	$\ln\Gamma(\eta+1) + (\eta+1)\ln 2$
베타	$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$	$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$	$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$	$\begin{bmatrix} \ln x \\ \ln(1-x) \end{bmatrix}$	$\ln \mathrm{B}(\alpha,\beta) = \ln\dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$
정규 기지 σ^2	μ	$\dfrac{\mu}{\sigma^2}$	$\sigma^2\eta$	x	$\dfrac{1}{2}\sigma^2\eta^2$
정규 미지 σ^2	$\begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$	$\begin{bmatrix} \dfrac{\mu}{\sigma^2} \\ -\dfrac{1}{2\sigma^2} \end{bmatrix}$	$\begin{bmatrix} -\dfrac{\eta_1}{2\eta_2} \\ -\dfrac{1}{2\eta_2} \end{bmatrix}$	$\begin{bmatrix} x \\ x^2 \end{bmatrix}$	$-\dfrac{\eta_1^2}{4\eta_2} - \dfrac{1}{2}\ln(-2\eta_2)$
다변량 정규	$\begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\Sigma} \end{bmatrix}$	$\begin{bmatrix} \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ -\dfrac{1}{2}\boldsymbol{\Sigma}^{-1} \end{bmatrix}$	$\begin{bmatrix} -\dfrac{1}{2}\boldsymbol{\eta}_2^{-1}\boldsymbol{\eta}_1 \\ -\dfrac{1}{2}\boldsymbol{\eta}_2^{-1} \end{bmatrix}$	$\begin{bmatrix} \mathbf{x} \\ \mathbf{x}\mathbf{x}^T \end{bmatrix}$	$-\dfrac{1}{4}\boldsymbol{\eta}_1^T\boldsymbol{\eta}_2^{-1}\boldsymbol{\eta}_1 - \dfrac{1}{2}\ln\lvert -2\boldsymbol{\eta}_2 \rvert$

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

표본 / 통계량

* 표본 → 1) 분포 f(x), p(x)의 추론 // 2) θ 추론 ← f(x), p(x)는 알고 있음 (X_i: 확률변수, x_i: 실현값)

1. 확률 표본 (Random sample): iid $[X_1, \cdots, X_n]$
2. 통계량 (Statistic): $T = T(X_1, \cdots, X_n)$ (표본에 대한 함수)

   ➔ ***$\theta \in \Omega$에 대한 추정량***이면 T: 점추정량 (point estimator), 실현값 t: 점추정값 (point estimate)
3. **불편추정량 (Unbiased estimator)**: $E(T) = \theta$ $\quad [E(\bar{X}) = \mu, \ \ E(S^2) = \sigma^2]$
4. **Maximum likelihood estimator (mle)**

 1) **가능도 함수:** $L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$ ← **MLE:** $\hat{\theta} = \mathrm{Argmax}[L(\theta)]$

 2) **로그우도 함수:** $l(\theta) = \sum_{i=1}^{n} \ln f(x_i, \theta)$ ← **MLE:** $\hat{\theta} = \mathrm{Argmax}[l(\theta)] \leftarrow \partial l / \partial\theta = 0$

지수	$l(\beta) = \sum_{i=1}^{n} \ln \frac{1}{\beta} e^{-\frac{x_i}{\beta}} = -\frac{1}{\beta}\sum x_i - n\ln\beta = -\mathrm{n}\left(\frac{1}{\beta}\bar{X} + \ln\beta\right) \to \frac{\partial l}{\partial\beta} = \mathrm{n}\left(\frac{\bar{\mathrm{X}}}{\beta^2} - \frac{1}{\beta}\right) \to \hat{\beta} = \bar{X}$ (also 불편)
이항	$l(p) = \sum_{i=1}^{n} \ln p^{x_i}(1-p)^{1-x_i} = n\bar{X}\ln p + (n - n\bar{X})\ln(1-p) \to \frac{\partial l}{\partial p} = n\left(\frac{\bar{\mathrm{X}}}{p} - \frac{1-\bar{X}}{1-p}\right) \to \hat{p} = \bar{X}$ (also 불편)
정규	$l(\mu,\sigma) = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2}\sum\left(\frac{x_i-\mu}{\sigma}\right)^2 \to \nabla l(\mu,\sigma) = \left[\frac{1}{\sigma}\sum\left(\frac{x_i-\mu}{\sigma}\right), -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum(x_i-\mu)^2\right]^T$ $\to \hat{\mu} = \bar{X}, \ \ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{\mathrm{n}}(X_i - \bar{X})^2 = \frac{n-1}{n}S^2$

CLT

1) Pivot 확률변수: (추정량-모수)/표준오차

2) 중심극한정리: $Z_n = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$ ($X_i \overset{iid}{\sim} \mathrm{random}(\mu,\sigma^2) \Rightarrow \bar{X} \xrightarrow{D} N\left(\mu, \frac{\sigma^2}{n}\right)$) cf) $X \sim N(\mu,\sigma^2)$이면 정확 정규분포

신뢰 구간

*신뢰구간: **모수 $\boldsymbol{\theta}$가 추정량 $\hat{\boldsymbol{\theta}}$에서** 얼마나 벗어났는가?

1. 신뢰구간: $1-\alpha = P_\theta[\theta \in (\hat{\theta}_L, \hat{\theta}_U)]$ → 100(1-α)% 신뢰구간 (같은 신뢰계수 → 구간 길이 최소화)

*해석: **모수 $\boldsymbol{\theta}$가 추정량 $(\hat{\boldsymbol{\theta}}_L, \hat{\boldsymbol{\theta}}_U)$ 구간에 있는 사건 $\sim B(1, 1-\alpha)$ (95% CI: $\boldsymbol{\theta}$가 $(\hat{\boldsymbol{\theta}}_L, \hat{\boldsymbol{\theta}}_U)$에 평균 19회/20회)**

2. 평균 신뢰 구간 ($\boldsymbol{z}_{\alpha/2}$: 상위 $\alpha/2$에서 z값) ⇔ $(z_{\alpha/2} = \xi_{1-\alpha/2} \leftrightarrow F(z_{\alpha/2}) = F(\xi_{1-\alpha/2}) = 1-\alpha/2)$

상황	가정	Pivot statistic	$\boldsymbol{\mu}$ 의 100(1-α)% CI
Z	평균 μ **분산 $\boldsymbol{\sigma^2}$**	$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$	$1-\alpha \approx P\left(-\boldsymbol{z}_{\alpha/2} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < \boldsymbol{z}_{\alpha/2}\right)$
T	$X_i \sim N(\mu,\sigma^2)$	$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ $(df = n-1)$	$1-\alpha = P\left(-\boldsymbol{t}_{\alpha/2,n-1} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < \boldsymbol{t}_{\alpha/2,n-1}\right)$

* 전부 N≥30 or 정규성 가정... (Z는 CLT로 근사 가능 / T는 robustness로 비정규 상황 적용 가능)

3. 평균 차이 $(\bar{\boldsymbol{X}} - \bar{\boldsymbol{Y}})$ 신뢰 구간 $* E(\bar{X}-\bar{Y}) = \mu_1 - \mu_2, \ \ \mathrm{Var}(\bar{X}-\bar{Y}) = (\sigma_1^2/n_1) + (\sigma_2^2/n_2)$

상황	가정	Pivot statistic	유도/비고
Z	평균 $\mu_1 - \mu_2$ 분산 $(\sigma_1^2/n_1) + (\sigma_2^2/n_2)$	$\boldsymbol{Z} = \frac{(\bar{\boldsymbol{X}}-\bar{\boldsymbol{Y}}) - (\boldsymbol{\mu_1}-\boldsymbol{\mu_2})}{\sqrt{(\boldsymbol{\sigma_1^2}/\boldsymbol{n_1}) + (\boldsymbol{\sigma_2^2}/\boldsymbol{n_2})}} \sim N(0,1)$	
T (등분산)	$X_i \sim N(\mu_1, \boldsymbol{\sigma^2})$ $Y_i \sim N(\mu_2, \boldsymbol{\sigma^2})$	$\boldsymbol{T} = \frac{(\bar{\boldsymbol{X}}-\bar{\boldsymbol{Y}}) - (\boldsymbol{\mu_1}-\boldsymbol{\mu_2})}{\boldsymbol{S_p}\sqrt{(\boldsymbol{1}/\boldsymbol{n_1}) + (\boldsymbol{1}/\boldsymbol{n_2})}}$ $(df = n_1 + n_2 - 2)$ $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1-1)+(n_2-1)}$	1) $\mathrm{E}(S_p^2) = \sigma^2$ 2) $(\mathrm{n}_1 - 1)S_1^2/\sigma^2 \sim \chi^2(n_1-1)$ $\to (n_1+n_2-2)S_p^2 \sim \chi^2(n_1+n_2-2)$ 3) $S_p^2 \leftrightarrow (\bar{X}-\bar{Y})$ 독립
T (이분산)	$X_i \sim N(\mu_1, \boldsymbol{\sigma_1^2})$ $Y_i \sim N(\mu_2, \boldsymbol{\sigma_2^2})$	$\boldsymbol{T} = \frac{(\bar{\boldsymbol{X}}-\bar{\boldsymbol{Y}}) - (\boldsymbol{\mu_1}-\boldsymbol{\mu_2})}{\sqrt{(\boldsymbol{S_1^2}/\boldsymbol{n_1}) + (\boldsymbol{S_2^2}/\boldsymbol{n_2})}}$ **(Welch's, df<-round(df))**	$df = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^2}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{n_x-1} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{n_y-1}}$
Paired T	$\boldsymbol{d_i} = x_i - y_i$ $\sim N(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)$	$\boldsymbol{T} = \frac{\bar{\boldsymbol{d}} - \boldsymbol{\mu_D}}{\sqrt{(\boldsymbol{S_1^2}+\boldsymbol{S_2^2})/\boldsymbol{n}}}$	$\mu_D = \mu_1 - \mu_2$

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

신뢰구간

4. 모비율 (극한 표준정규분포; CLT) *정규 근사 시 연속성 수정 적용 가능

상황	가정	Exact binomial	정규근사
단일 모비율	$X \sim B(n,p) \xrightarrow{d} N(np, np(1-p))$ **표본 비율** $\hat{p} = \frac{X}{n} \xrightarrow{d} N\left(p, \frac{p(1-p)}{n}\right)$	① $H_1: p \neq p_0$ P-값=2 $\mathbf{min}\{P_{H_0}(\hat{\boldsymbol{p}} > \boldsymbol{p_0}), P_{H_0}(\hat{\boldsymbol{p}} < \boldsymbol{p_0})\}$ ② $H_1: p > p_0$ P-값=$P_{H_0}(\hat{\boldsymbol{p}} > \boldsymbol{p_0})$	$\boldsymbol{Z} = \dfrac{\hat{\boldsymbol{p}} - \boldsymbol{p}}{\sqrt{\dfrac{\hat{\boldsymbol{p}}(\boldsymbol{1}-\hat{\boldsymbol{p}})}{\boldsymbol{n}}}}$
모비율 차이	표본 비율 차이: $\hat{p}_1 - \hat{p}_2$ ① 평균 $p_1 - p_2$ ② 분산 $\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$		$\boldsymbol{Z} = \dfrac{(\hat{\boldsymbol{p}}_1 - \hat{\boldsymbol{p}}_2) - (\boldsymbol{p}_1 - \boldsymbol{p}_2)}{\sqrt{\dfrac{\hat{\boldsymbol{p}}_1(\boldsymbol{1}-\hat{\boldsymbol{p}}_1)}{\boldsymbol{n}_1} + \dfrac{\hat{\boldsymbol{p}}_2(\boldsymbol{1}-\hat{\boldsymbol{p}}_2)}{\boldsymbol{n}_2}}}$

5. 이산형 모수

1) $F_T(T;\theta)$: 통계량 T의 cdf; θ에 대해 단조 감소 → 신뢰 구간: $F_T(T_{n-1};\ \underline{\theta}) = 1-\alpha_2,\ F_T(T_n;\bar{\theta}) = \alpha_1$

2) **Bisection algorithm**: 순감소 $g(x) = d \in g([a,\ b])$ → 1) if $g\{(a+b)/2\} > d$ → 구간 $[(a+b)/2,\ b]$ 재설정
→ 2) if $g\{(a+b)/2\} < d$ → 구간 $[a,\ (a+b)/2]$ 재설정

상황	조건	유도
Binomial	$X \sim b(1,p)$ $n = 30, \bar{x} = 0.60$ $T = n\bar{X} \sim b(30,p)$ $(T_{n-1} = 17,\quad T_n = 18)$	① **하한:** pbinom(17, 30, 0.4)=0.9787, pbinom(17,30,0.45)=0.9286 ➔ pbinom(17, 30, **0.434**) ≈ 0.95 ② **상한:** pbinom(18, 30, 0.7)=0.1593, pbinom(18,30,0.8)=0.0094 ➔ pbinom(18, 30, **0.747**) ≈ 0.05 ∴ $\boldsymbol{p}$의 90% CI: $[\mathbf{0.434}, \mathbf{0.747}]$
Poisson	$X \sim \mathrm{Poi}(\mu)$ $\mathrm{n} = 25, \bar{x} = 5$ $T = n\bar{X} \sim \mathrm{Poi}(25\mu)$ $(T_{n-1} = 124,\quad T_n = 125)$	① **하한:** ppois(124, 25 x 4)=0.9912, ppois(124, 25 x 4.4)=0.9145 ➔ ppois(124, 25 x **4.287**) ≈ 0.95 ② **상한:** ppois(125, 25 x 5.5)=0.1330, ppois(125, 25 x 6)=0.0204 ➔ ppois(125, 25 x **5.8**) ≈ 0.05 ∴ $\boldsymbol{\mu}$의 90% CI: $[\mathbf{4.287}, \mathbf{5.8}]$

순서통계량

*정의: $(Y_1 < \cdots < Y_n) \leftarrow [X_1, \cdots, X_n]$ 재배열 *강점: 분포에 종속되지 않음.

1. **PDF**: $g(y_1, \cdots, y_n) = \boldsymbol{n}!\, f(y_1) \cdots f(y_n)$ (on $a < y_1 < \cdots < y_n < b$) pf) $g(y_1, \cdots, y_n) = \sum_{i=1}^{n!} |J_i| f(y_1) \cdots f(y_n)$

2. **Marginal PDF** 1) $\boldsymbol{g_k(y_k)} = \dfrac{\boldsymbol{n}!}{(\boldsymbol{k}-\boldsymbol{1})!\,(\boldsymbol{n}-\boldsymbol{k})!\,(\boldsymbol{1})!} [\boldsymbol{F}(\boldsymbol{y_k})]^{\boldsymbol{k}-\boldsymbol{1}} [\boldsymbol{1} - \boldsymbol{F}(\boldsymbol{y_k})]^{\boldsymbol{n}-\boldsymbol{k}} \boldsymbol{f}(\boldsymbol{y_k})$

pf) $g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} \int_{y_k}^b \cdots \int_{y_{n-1}}^b \boldsymbol{n}!\, f(y_1) \cdots f(y_n)\, dy_n \cdots dy_{k+1} dy_1 \cdots dy_{k-1}$ $(y_n \to y_{k+1}\,; y_1 \to y_{k-1})$

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

3. **Quantile (분위수)**: cdf $F(\xi_p) = p \leftrightarrow \boldsymbol{\xi_p} = \boldsymbol{F^{-1}}(\boldsymbol{p})$, $\boldsymbol{k} = \mathbf{floor}[\boldsymbol{p}(\boldsymbol{n}+\boldsymbol{1})]$

1) $\boldsymbol{F}(\boldsymbol{Y_k})$는 $\dfrac{\boldsymbol{k}}{\boldsymbol{n}+\boldsymbol{1}}$의 불편추정량: $E(F(Y_k)) = \int_a^b F(y_k) g_k(y_k)\, dy_k = \int_0^1 \dfrac{n!}{(k-1)!\,(n-k)!} z^k (1-z)^{n-k} dz = \dfrac{k}{n+1}$

2) Quartile: 1분위수 $(\boldsymbol{Q_1} = Y_{[0.25(n+1)]})$ ⬄ 중위수 $(\boldsymbol{Q_2} = Y_{[0.5(n+1)]})$ ⬄ 3분위수 $(\boldsymbol{Q_3} = Y_{[0.75(n+1)]})$

*중위수: 홀수➔중간값 $Y_{(n+1)/2}$ / 짝수➔ $(Y_{(n/2)} + Y_{(n/2)+1})/2$

➔ **Box plot**: $h = 1.5(Q_3 - Q_1)$, $LF = Q_1 - h$, $UF = Q_3 + h$ (LF, UF 바깥: 이상값; 정규분포상 P≤0.007)

3) **Q-Q plot**: 표본의 순서통계량 $(Y_1, Y_2, \cdots, Y_{50})$ ⬄ 이론적 분위수 $(Z_{0.02}, Z_{0.04}, \cdots, Z_{1.00})$ ← any 분포

4) 신뢰구간: $1 - \alpha = P(Y_i < \xi_p < Y_j) = \sum_{w=i}^{j-1} \binom{n}{w} p^w (1-p)^{n-w}$ ← $\boldsymbol{p} = \boldsymbol{F}(\boldsymbol{\xi_p})$ (중위수: $p = 1/2$)

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

가설 검정

1) 가설 정의: $H_0: \theta \in \boldsymbol{\omega_0}$ (Null) vs. $H_1: \theta \in \boldsymbol{\omega_1}$ (alternative) ← $X \sim f(x;\theta)$에 대해 $\theta \in \Omega = (\omega_0 \cup \omega_1)$, 분할

2) 가설 검정: 표본 $(X_1, \cdots, X_n) \in C$ → H_1채택 (기각역 $C \subset D = \mathrm{span}\{(X_1, \cdots, X_n)\}$)

표본 $(X_1, \cdots, X_n) \notin C$ → H_0유지

3) 유의 수준: $\alpha = \max_{\theta \in \omega_0} P_\theta[(X_1, \cdots, X_n) \in C]$ (복합귀무가설에 대해 모든 null 모수 → 기각역에 속할 확률 최대)

* 1종 오류: H_0 참, but 기각 → H_1 채택 (=FP) ∴ 유의수준(α): 1종 오류 범할 최대 확률

4) 검정력: $\gamma(\theta) = P_\theta[(X_1, \cdots, X_n) \in C],\ \theta \in \omega_1$ → 유의 수준과 다르게 대립가설 모수에 따라 달라짐

① 2종 오류: H_0 거짓, but 유지 → H_0 유지 (=FN) ∴ β: 2종 오류 범할 확률 (under given $\theta \in \boldsymbol{\omega_1}$)

② 검정력: H_0 거짓 → 알맞게 H_1 채택 (TP)

5) **P-값**: 1) **Upper tail**: **P-값**$= \boldsymbol{P_{H_0}(X \geq x_{obs}) = 1 - F_{H_0}(x_{obs})}$

2) **Lower tail**: **P-값**$= \boldsymbol{P_{H_0}(X \leq x_{obs}) = F_{H_0}(x_{obs})}$

3) **2-sided**: **P-값**$= \boldsymbol{2 \times P_{H_0}(X \geq |x_{obs}|) = 2[1 - F_{H_0}(|x_{obs}|)]}$ (X=0 좌우 대칭)

→ $X = F^{-1}(U)$ (단조 증가) 정리의 역에 의해 P-값 ~ unif(0,1) under 귀무가설 H_0

예시	분포	가설	유도
단일 이항 단측	$X_i \sim B(1,p)$	$H_0: p = p_0$ $H_1: p < p_0$	*표본통계량: $\boldsymbol{S = \sum_{i=1}^n X_i \sim B(n,p)}$ 1) 기각역 설정: 귀무가설 하에서 $S \sim B(n,p_0) \to \alpha = P_{p_0}[S \leq k]$ → $0.11 = P_{p_0}[S \leq 11]$ $(n = 20,\ p_0 = 0.7)$ 2) 검정력 함수: $\gamma(p) = P_p[S \leq 11]$ (**단조 감소 of p**) ∴ <u>$\boldsymbol{H_0: p \geq p_0}$**로 확장**</u> ← $\max_{p \geq p_0} P_p[S \leq k] = P_{p_0}[S \leq k]$ (**단조성**)
	대표본에서 $\boldsymbol{\dfrac{\hat{p} - p_0}{\sqrt{\hat{p}(1-\hat{p})/n}} \approx \dfrac{\hat{p} - p_0}{\sqrt{p(1-p)/n}} \xrightarrow{D} N(0,1)}$		
대표본			* 표본통계량: $\boldsymbol{\dfrac{\bar{X} - \mu}{S/\sqrt{n}} \approx \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)}$
대표본 단측 (Upper)	X_i ~ **미지** 분포 1) 평균: μ 2) 분산: σ^2	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	1) 기각역 설정: $\alpha = P_{\mu_0}\left[\dfrac{\boldsymbol{\bar{X} - \mu_0}}{\boldsymbol{S/\sqrt{n}}} \geq z_\alpha\right] \approx 1 - \Phi(z_\alpha)$ 2) 검정력 함수: $\gamma(\mu) = P_\mu\left[\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right] = P_\mu\left[\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \dfrac{\mu_0 - \mu}{\sigma/\sqrt{n}} + z_\alpha\right]$ $\approx 1 - \Phi\left(\dfrac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + z_\alpha\right) = \boldsymbol{\Phi\left(\dfrac{\sqrt{n}(\mu - \mu_0)}{\sigma} - z_\alpha\right)}$ (**단조 증가 of μ**) * **Power 증가: n↑, 효과크기 (μ-μ₀)↑, α↑ & σ↓**
대표본 단측 (Lower)		$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	1) 기각역 설정: $\alpha = P_{\mu_0}\left[\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -z_\alpha\right] \approx \Phi(z_\alpha)$ 2) 검정력 함수: $\gamma(\mu) = P_\mu\left[\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha\right] = P_\mu\left[\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \dfrac{\mu_0 - \mu}{\sigma/\sqrt{n}} - z_\alpha\right]$ $\approx \boldsymbol{\Phi\left(\dfrac{\sqrt{n}(\mu_0 - \mu)}{\sigma} - z_\alpha\right) = \Phi\left(-\dfrac{\sqrt{n}(\mu - \mu_0)}{\sigma} - z_\alpha\right)}$ (**단조 감소 of μ**) * **Power 증가: n↑, 효과크기 (μ-μ₀)↑, α↑ & S↓**
대표본 양측		$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	1) 기각역 설정: $\boldsymbol{\alpha} = P_{\mu_0}\left[\left\lvert\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}\right\rvert \geq z_{\alpha/2}\right]$ ⇐ (**양측 동일 배분**) 2) 검정력 함수: $\gamma(\mu) = P_\mu\left[\left\lvert\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right\rvert \geq z_{\alpha/2}\right]$ $\approx \boldsymbol{\Phi\left(\dfrac{\sqrt{n}(\mu - \mu_0)}{\sigma} - z_{\alpha/2}\right) + \Phi\left(-\dfrac{\sqrt{n}(\mu - \mu_0)}{\sigma} - z_{\alpha/2}\right)}$ (**U자 함수 of μ**) → (μ₀에서 최소값)
t-검정 정규성	$X_i \sim N(\mu, \sigma^2)$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	* 표본통계량: $\boldsymbol{T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)}$ * **t-분포**는 N(0,1) 보다 누워 있음 → "보수적" // 정규성 하 "정확"
2-표본 t-검정	$X_i \sim N(\mu_1, \boldsymbol{\sigma^2})$ $Y_i \sim N(\mu_2, \boldsymbol{\sigma^2})$ (정규,등분산)	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	* 표본통계량: $\boldsymbol{T = \dfrac{(\bar{X} - \bar{Y}) - 0}{S_p\sqrt{(1/n_1) + (1/n_2)}} \sim t(n_1 + n_2 - 2)}$ * $\boldsymbol{\lvert T\rvert \geq t_{0.025, n_1+n_2-2}}$ 이면 $\boldsymbol{H_0}$ **기각**

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

Pearson χ² 검정

자유도: n(확률표본)-n(미지수 or 제약)

2 cells

1. 상황: $\boldsymbol{X_1 \sim b(n, p_1)}$, $X_2 = n - X_1$, $p_2 = 1 - p_1 \;\Rightarrow\; Y = \dfrac{X_1 - np_1}{\sqrt{np_1(1-p_1)}} \xrightarrow{D} N(0,1)$; $Q_1 = Y^2 \xrightarrow{D} \chi^2(1)$

2. 검정통계량: $\boldsymbol{Q_1} = \dfrac{(X_1 - np_1)^2}{np_1(1-p_1)} = \dfrac{(X_1 - np_1)^2}{np_1} + \dfrac{(X_1 - np_1)^2}{n(1-p_1)} = \dfrac{\boldsymbol{(X_1 - np_1)^2}}{\boldsymbol{np_1}} + \dfrac{\boldsymbol{(X_2 - np_2)^2}}{\boldsymbol{np_2}} \xrightarrow{D} \boldsymbol{\chi^2(1)}$

k cells

1. 상황: k항; n회 다항분포 $(p_k = 1 - \sum_{i=1}^{k-1} p_i$ & $x_k = n - \sum_{i=1}^{k-1} x_i)$

2. 검정통계량: $\boldsymbol{Q_{k-1}} = \displaystyle\sum_{i=1}^{k} \frac{\boldsymbol{(X_i - np_i)^2}}{\boldsymbol{np_i}} \xrightarrow{D} \boldsymbol{\chi^2(k-1)}$ $\Leftarrow (k-1)$개 알면 나머지 1개 앎

적합도 (GoF) 검정

1) 귀무가설: $H_0: p_1 = p_{1,0}$, $p_2 = p_{2,0}$, $\cdots$, $p_k = p_{k,0}$ ex) 분할표 검정 / **분포 검정** (구간별)

2) 검정통계량 ① 피어슨: $Q_{k-1} = \displaystyle\sum_{i=1}^{k} \frac{\boldsymbol{(X_i - E_i)^2}}{\boldsymbol{E_i}} = \sum_{i=1}^{k} \frac{\left(X_i - np_{i,0}\right)^2}{np_{i,0}} \xrightarrow{D} \boldsymbol{\chi^2(k-1)}$ (귀무가설 하)

② 로그우도비: $G^2 = 2\displaystyle\sum_{i=1}^{k} x_i \ln\left(\frac{x_i}{e_i}\right) \rightarrow \boldsymbol{P\text{값} = P\left(\chi^2_{k-1} \geq Q_{obs}\right)}$ $*$ Upper tail 단측

*로그우도비 증명: $L(\boldsymbol{\theta_0}) = \frac{n!}{\prod(x_i)!} \prod p_{i,0}^{x_i}$, $L(\hat{\boldsymbol{\theta}}) = \frac{n!}{\prod(x_i)!} \prod \left(\frac{x_i}{n}\right)^{x_i}$ ➔ $-2\ln\Lambda = 2\sum_{i=1}^{k} x_i \ln\left(\frac{x_i}{e_i}\right) \xrightarrow{d} \chi^2_{k-1}$

최소 χ² 추정량

<예시> 정규분포 모수 추정 $N(\mu, \sigma^2)$

1) 상황: 실수구간 → k등분 $(A_1, \cdots, A_k)$; $\boldsymbol{p_i} = \int_{\boldsymbol{A_i}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}(y - \boldsymbol{\mu})^2/\boldsymbol{\sigma}^2\right] dy$

2) 실제 A_i의 도수인 X_i → $Q_{k-1} = \sum_{i=1}^{k} \frac{(X_i - np_i)^2}{np_i} \xrightarrow{D} \boldsymbol{\chi^2(k-3)}$ **최소화하는** $\boldsymbol{\hat{\mu}, \hat{\sigma}^2}$

r x c cells

동질성 검정

1) 상황: **2개의 k항 다항분포** *각 모수: $(\boldsymbol{n_1}, p_{11}, p_{21}, \cdots, p_{k1}), (\boldsymbol{n_2}, p_{12}, p_{22}, \cdots, p_{k2})$

$\Rightarrow \displaystyle\sum_{j=1}^{2}\sum_{i=1}^{k} \frac{\left(X_{ij} - n_j p_{ij}\right)^2}{n_j p_{ij}} \xrightarrow{D} [\chi^2(k-1) + \chi^2(k-1)] = \chi^2(2k-2)$

2) 귀무가설: $H_0: p_{11} = p_{12}$, $p_{21} = p_{22}$, $\cdots$, $p_{k1} = p_{k2}$ (둘은 구간 별 비율이 동일)

$\Rightarrow p_{m1} = p_{m2}$의 MLE: $\dfrac{X_{m1} + X_{m2}}{n_{m1} + n_{m2}}$ (총 $k-1$개 점추정값 필요)

3) 검정통계량: $\displaystyle\sum_{j=1}^{2}\sum_{i=1}^{k} \frac{\left[\boldsymbol{X_{ij} - n_j\left(\frac{X_{i1} + X_{i2}}{n_{i1} + n_{i2}}\right)}\right]^2}{\boldsymbol{n_j\left(\frac{X_{i1} + X_{i2}}{n_{i1} + n_{i2}}\right)}} \xrightarrow{D} \boldsymbol{\chi^2(k-1)}$ (귀무가설 하)

독립성 검정

1) 상황: 확률실험 n회 결과 → 가로 (A) a항 / 세로 (B) b항 **두 종류 범주**로 구분

➔ $p_{ij} = P\left(A_i \cap B_j\right)$, X_{ij}는 $A_i \cap B_j$ 도수

$\Rightarrow Q_{ab-1} = \displaystyle\sum_{j=1}^{b}\sum_{i=1}^{a} \frac{\left(X_{ij} - np_{ij}\right)^2}{np_{ij}} \xrightarrow{D} \chi^2(ab-1)$

2) 귀무가설: $H_0: P\left(A_i \cap B_j\right) = P(A_i)P\left(B_j\right)$ for all (i,j) (속성 A, B는 독립)

$\Rightarrow p_{i*} = P(A_i)$의 MLE: $\hat{p}_{i*} = \dfrac{X_{i*}}{n} = \dfrac{\sum_{j=1}^{b} X_{ij}}{n}$ [총 (a − 1) + (b − 1)개 점추정값 필요]

3) 검정통계량: $\displaystyle\sum_{j=1}^{b}\sum_{i=1}^{a} \frac{\left[\boldsymbol{X_{ij} - n\left(\frac{X_{i*}}{n}\right)\left(\frac{X_{*j}}{n}\right)}\right]^2}{\boldsymbol{n\left(\frac{X_{i*}}{n}\right)\left(\frac{X_{*j}}{n}\right)}} \xrightarrow{D} \boldsymbol{\chi^2[(a-1)(b-1)]}$ (귀무가설 하)

비중심 분포

비중심 χ²

$\boldsymbol{Y} = \displaystyle\sum_{i=1}^{n} \frac{\boldsymbol{X_i^2}}{\boldsymbol{\sigma^2}}$, $X_i \sim N(\mu_i, \sigma^2)$ $*\,\mu_i = 0$이면 $Y \sim \chi^2(n)$

$M(t) = E[\exp(tX_i^2/\sigma^2)] = \dfrac{1}{(1-2t)^{n/2}} \exp\left[\dfrac{t\sum_{i=1}^{n}\mu_i^2}{\sigma^2(1-2t)}\right] = \dfrac{1}{(1-2t)^{n/2}} \exp\left[\dfrac{t}{1-2t}\theta\right]$ $\left(t < \dfrac{1}{2}\right)$

$\boldsymbol{Y} = \displaystyle\sum_{i=1}^{n} \frac{\boldsymbol{X_i^2}}{\boldsymbol{\sigma^2}} \sim \chi^2(n, \theta)$ $\left(\theta = \dfrac{\sum_{i=1}^{n}\mu_i^2}{\sigma^2}\right)$ $pf)$ MGF 적분 활용 → 치환하여 정규분포 PDF꼴로 정리

* R codes 1) dchisq (x,r,a): f(X=x) 2) pchisq (x,r,a): P(X≤x)

비중심 F

$U \sim \chi^2(n_1, \theta)$ & $V \sim \chi^2(n_2)$ $*\,U, V$는 독립

$Y = \dfrac{U/n_1}{V/n_2} \sim F(n_1, n_2, \theta)$

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

Monte Carlo

*** 몬테카를로 생성: 특정 "Known" 표본/분포 ➔ 관측값 생성** (Resampling, Bayesian 등에서 중요)

1. 균등분포 (Uniform distribution): $\text{unif}(a,b)$; $\text{pdf} = 1/(\text{b}-\text{a})$ ← R codes: runif(횟수)

unif(0,1)⇔CDF "관측치 생성"

$\boldsymbol{X = F^{-1}(U)}$는 $\mathbf{cdf}\ \boldsymbol{F(X)}$ 따름⇔ 역: $Z = F(X) \sim \text{unif}(0,1)$

지수	$F(x) = 1 - e^{-x/\beta}, \quad (x > 0)$ $\therefore X = F^{-1}(U) = -\beta \ln(1-U)$ 는 지수분포 생성
푸아송	$m = \lambda w$ → $T_i \sim \exp(1/\lambda)$ 에 대해 $[X = k] \Leftrightarrow \sum_{i=1}^{k} T_i \le w$ & $\sum_{i=1}^{k+1} T_i > w$
	* 구간 w 동안 난수로 T_i 생성 → 횟수 카운트 (초기 $X = 0, T = 0$) 1) $\Delta T = -(1/\lambda)\ln(1-U)$ 2) $T \leftarrow T + \Delta T$ 3) if $\text{T} \le \text{w}$: $X \leftarrow X + 1$ elif $\text{T} > \text{w}$: return X
정규 분포	<**Box & Muller** (1958)> → 일반화: Marsaglia & Bray (1964) $X_1 = (-2\ln Y_1)^{1/2} \cos(2\pi Y_2)$; $X_2 = (-2\ln Y_1)^{1/2} \sin(2\pi Y_2) \Leftarrow Y_1, Y_2 \sim \text{unif}(0,1)$ $f(X_1, X_2) = \lvert J \rvert g(Y_1, Y_2) = \dfrac{1}{2\pi} \exp\left[-\dfrac{1}{2}(x_1^2 + x_2^2)\right]$

채택-기각 (A-R) 알고리즘 (어려운 CDF)

$X = F^{-1}(U)$를 closed form 계산 불가.⇐ $g(x)$ 이용: ①Easy ②$f(x)$ 유사 ③$\dfrac{f(x)}{g(x)} \le k$ (유계)

① $Y \sim g(y)$ & U 생성

② $U \le \dfrac{f(Y)}{kg(Y)} \le 1$이면 $\text{X} = \text{Y}$, 아니면 ①로 돌아가 재 생성 ⇒ 조금 더 넓은 $kg(x)$로 근사

($f(x) = cf_1(x)$와 $g(x) = dg_1(x)$적당히 상수배 하여 k 무시 가능)

감마 CDF $\boldsymbol{\Gamma(\alpha,\beta)}$	$Y_i \sim \Gamma(1,1)$ ➔ $X = \sum_{i=1}^{\alpha} Y_i \sim \Gamma(\alpha, 1)$ (α 정수: CDF 생성 쉬움) $X \sim \Gamma(\alpha, 1)$ ➔$\beta X \sim \Gamma(\alpha, \beta)$ (α 실수 → **문제!**)
	① $X \sim \Gamma(\alpha, 1)$ & $Y \sim \Gamma([\alpha], 1/b)$ ② $\dfrac{f(x)}{g(x)} = b^{-[\alpha]} x^{\alpha-[\alpha]} e^{-(1-b)x} \le b^{-[\alpha]} \left\{\dfrac{\alpha - [\alpha]}{(1-b)e}\right\}^{\alpha-[\alpha]}$ (by x로 미분) ③ 위 식을 b로 미분 ⇒ $\dfrac{f(x)}{g(x)} \le ([\alpha]/\alpha)^{-[\alpha]} \left\{\dfrac{\alpha - [\alpha]}{(1-[\alpha]/\alpha)e}\right\}^{\alpha-[\alpha]} = M$
정규 CDF $\boldsymbol{N(0,1)}$	① **Y~Cauchy (역 CDF 알려짐) ➔ X~N(0,1)** ② $\dfrac{\phi(x)}{g(x)} = \dfrac{(2\pi)^{-1/2}\exp\{-x^2/2\}}{\{\pi(1+x^2)\}^{-1}} \Rightarrow \dfrac{\phi_1(x)}{g_1(x)} = (1+x^2)\exp\{-x^2/2\} \le 2e^{-1/2} = M$

Monte Carlo t-검정 (오염된 정규)

$W \sim N(0, 1^2)$ & $W \sim N(0, \sigma_c^2)$ (ε: 0.25, $\sigma_c = 25$) ← $W = Z \text{ or } \sigma_c Z$; $E(W) = 0$

* 가설: $H_0: \mu = 0$, $H_1: \mu > 0$ 1) $\text{n} = 20$, 2) $t_{0.05,19} = 1.729$	* 추정 알고리즘 (N: 시뮬레이션 수) 1) $\text{n} = 20$ 표본 생성 ⇐ X (오염 정규; μ 모름) 분포 2) $T = (\bar{X} - \mu)/(S/\sqrt{n})$계산 ➔ 1),2) N번 반복 3) **유의수준 실험적 추정량:** $\boldsymbol{\hat{\alpha} = I/N}$ ($I: T > t_{0.05,10}$ 도수) $\mathbf{SE} = \sqrt{\boldsymbol{\hat{\alpha}(1-\hat{\alpha})/N}}$ 예시) $\hat{\alpha} = 0.0412 \pm 0.0039$

Monte Carlo 적분

적분가능한 $g(x)$의 closed form 역도함수 (≈부정적분) 존재X ➔ **수치적 적분**

$$\int_a^b g(x)\,dx = (b-a)\int_a^b g(x)\left(\frac{1}{b-a}\right)dx = \boldsymbol{(b-a)E[g(X)]} \Leftarrow X \sim \text{unif}(a,b)$$

$\therefore \boldsymbol{\bar{Y}} = \dfrac{1}{n}\sum_{i=1}^{n} \boldsymbol{Y_i} = \dfrac{1}{n}\sum_{i=1}^{n} \boldsymbol{(b-a)g(X_i)}$ 는 정적분의 unbiased estimator ⇐ $X_i \sim \text{unif}(a,b)$

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

Boot-strap 기본

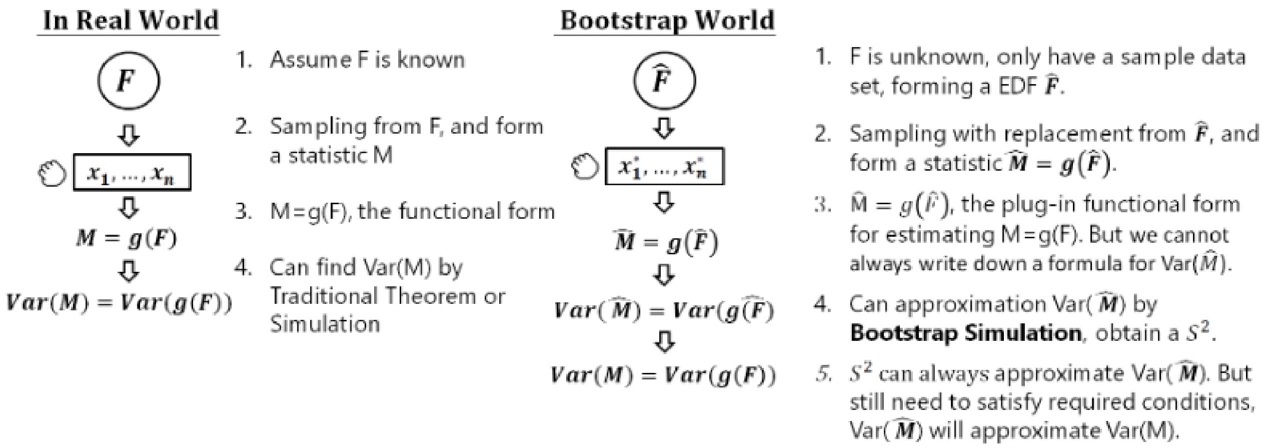
비교

1) 중심극한정리: 표본 통계량 $(\hat{\theta})$ 의 pivotal statistic이 극한 정규분포따름 ➔ 모수 θ 추정
2) 몬테카를로 기법: **X의 known 분포 (CDF)➔** 균등분포 난수추출기로 관측값 $X = F^{-1}(U)$ 생성
3) 부트스트랩: **X의 unknown 분포 ➔ 표본** $(X_1, \cdots X_n)$의 EDF $(\hat{F}_n)$→ 무작위 추출로 X_i^* 생성
$\hat{\theta}^*$의 분포 → $\hat{\theta}$의 신뢰구간 추정 → θ의 근사적 신뢰구간

원리

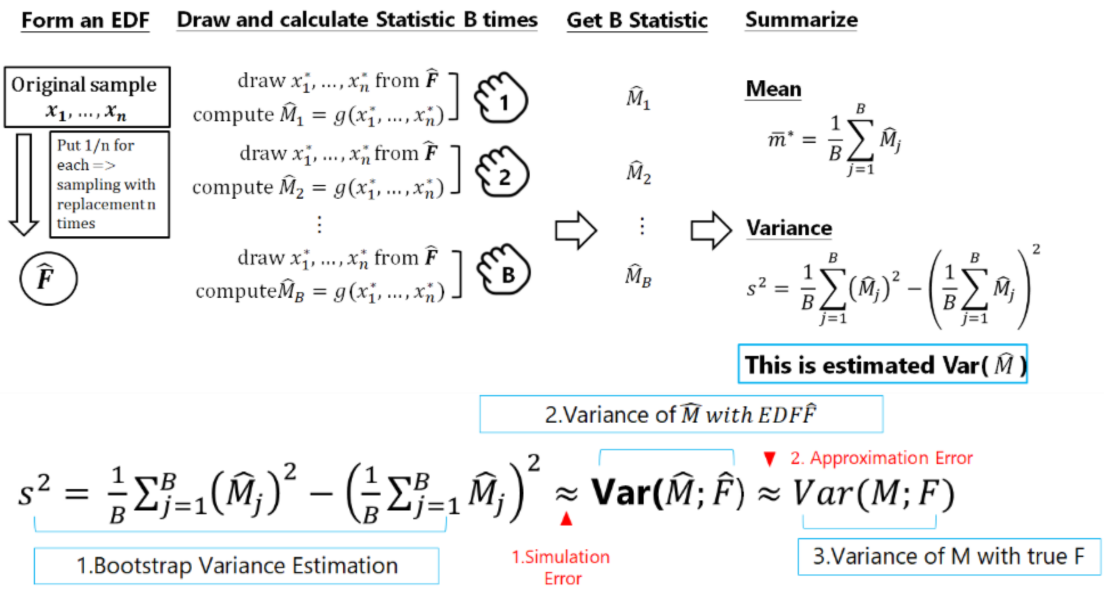
일반적인 통계적 추론에서는 estimator → parameter를 추정함.
Standard error는 estimator의 자체적인 변동성 (표준편차) (e.g. $SE(\bar{X}) = \sqrt{Var(\bar{X})} = \sigma/\sqrt{n}$)
Estimated SE는 SE에 unknown parameter가 들어가 있을 때, 다른 estimator를 이용 $(S/\sqrt{n})$

*문제: ①일반적인 확률변수 Y에 대해 분포 (PDF, CDF)를 알기 어렵고
② 통계량 g(Y)의 S.E.를 $\sigma/\sqrt{n}$처럼 정확한 수식으로 알아낼 수 있는 경우는 많지 않음.



EDF $\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \le x)$ ⇔ $\left(x\text{보다 작은 표본내 실현값 수}\right)$ / n ⇔ PMF는 $\frac{1}{n}$ (for every x_i)

Statistical functional (통계적 범함수): 모수가 [분포함수]의 함수로 표현됨. (평균,분산,중위수,백분위수,etc)
e.g. $E(X) = \int xfdx = \int xdF$, $Var(X) = \int x^2dF - \left(\int xdF\right)^2$
Plug-in principle: $\hat{\theta} = g(\hat{F}) = \int r(x)d\hat{F} \leftrightarrow \theta = g(F) = \int r(x)dF$ (전자는 후자의 plug-in estimate)



1번 simulation error는 결국 큰 수의 법칙에 의해 확률 수렴하므로 B↑으로 최소화 가능
2번 approximation error는 $\hat{F}$이 F에 근사 (n↑) 하면 최소화 (n↑면 자연스럽게 $\hat{M} \xrightarrow{P} M$ 성질도...)

7. 통계적 추론 – MLE / 신뢰구간 / 가설검정

Boot-strap 응용	**모 평균 추정**		$E(X_i^*) = \sum_{j=1}^{n}\frac{1}{n}X_j = \bar{X}$, $\mathrm{Var}(X_i^*) = \sum_{j=1}^{n}\frac{1}{n}(X_j-\bar{X})^2 = \frac{n}{n-1}S^2$ $E(\bar{X}_j^*) = \bar{X}$, $\mathrm{Var}(\bar{X}_j^*) = \frac{S^2}{n-1}$ B회 시뮬레이션 평균 $\frac{1}{B}\sum_{i=1}^{B}\bar{X}_j^* \xrightarrow{P} E(\bar{X}_j^*) = \bar{X} \xrightarrow{P} \boldsymbol{\mu}$, 분산 $\frac{1}{B}\sum_{i=1}^{B}(\bar{X}_j^*)^2 - \left(\frac{1}{B}\sum_{i=1}^{B}\bar{X}_j^*\right)^2 \xrightarrow{P} \mathrm{Var}(\bar{X}_j^*) = \frac{S^2}{n-1} \xrightarrow{P} \frac{\boldsymbol{\sigma^2}}{\boldsymbol{n}}$ ➔ B 회 부트스트랩 $\bar{X}$ 신뢰구간 (비모수적 counting) ≈ $\left[\bar{X} - z_{\alpha/2}\frac{S^2}{n},\ \bar{X} + z_{\alpha/2}\frac{S^2}{n}\right]$ ≈ [μ의 CLT 신뢰 구간] - 위의 정규가정을 통한 $z_{\alpha/2}$ 근사는 책 참고 4.9.1를 참조 - 다른 모수 추정도 크게 다르지 않음. ($\bar{X}$처럼 precise한 분산식이 존재하지 않으면 시뮬레이션 효과↑) - 통계량의 분포가 다른 모수에 종속되지 않게 pivot화하면 부트스트랩 정확성 향상 가능
	Boot-strap 검정	<1표본 평균> $H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	1. 상황: 1) 검정통계량: $\bar{X}$ 2) $\hat{p} = P_{H_0}[\bar{X} \ge \bar{x}]$ 2. H_0 가정 ⇒ $\boldsymbol{z_i = x_i - \bar{x} + \mu_0}$ $\left(E(z_i^*) = E(\bar{z}_j^*) = \mu_0\right)$ ⇒ 복원으로 n개 부트스트랩 3. **Empirical P-value 산출:** $\boldsymbol{\hat{p} = I/B}$ $(I: \{\bar{z}_j^* > \bar{x}\})$
		<2표본 평균> $H_0: \mu_2 = \mu_1$ $H_1: \mu_2 > \mu_1$	1. 상황: 1) 검정통계량: $V = \bar{Y} - \bar{X}$ 2) $\hat{p} = P_{H_0}[V \ge \bar{y} - \bar{x}]$ 2. H_0 가정 ⇒ 표본 합침 $(n = n_1 + n_2)$ ⇒ 복원으로 $\left(\mathbf{X}_\mathbf{i}^*, n_\mathbf{1}\text{개}\right), \left(\mathbf{Y}_\mathbf{i}^*, n_\mathbf{2}\text{개}\right)$ 추출 3. **Empirical P-value 산출:** $\boldsymbol{\hat{p} = I/B}$ $(I: \{\bar{y}_i^* - \bar{x}_i^* > \bar{y} - \bar{x}\})$ * 부연: H_0 가정 했기 때문에 생성값 $(\bar{y}_i^* - x_i^*)$은 H_0하 통계량임.
	Perm test		2표본 perm test: 통합 표본 ($n=n_1+n_2$)에서 비복원으로 추출된 x,y 모든 가능한 표본→ 검정

8. 정규모형 추론: ANOVA (*T-검정처럼 ANOVA 역시 관측치 분포에 robust함)

de1-Way ANOVA

	귀무가설	대립가설
전제	정규성+등분산성 $\boldsymbol{X_{ij} \sim N(\mu_i, \sigma^2)} \Leftrightarrow \boldsymbol{\epsilon_{ij} \sim N(0, \sigma^2)}$ $(x_{ij} = \mu_i + \epsilon_{ij})$	
$\boldsymbol{X_{ij}}$	$\boldsymbol{X_{ij} \overset{iid}{\sim} N(\mu, \sigma^2)}, \quad n_T = \sum_{i=1}^{k} n_i$	$\boldsymbol{X_{ij} \overset{iid}{\sim} N(\mu_i, \sigma^2)}$
표본평균	$\bar{X}_{\cdot\cdot} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}}{n_T} \sim N\left(\mu, \frac{\sigma^2}{n_T}\right)$ $\bar{X}_{i\cdot} = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i} \sim N\left(\mu, \frac{\sigma^2}{n_i}\right)$ $\bar{X}_{\cdot j} = \frac{\sum_{i=1}^{k} X_{ij}}{k} \sim N\left(\mu, \frac{\sigma^2}{\boldsymbol{k}}\right)$	$\bar{X}_{\cdot\cdot} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}}{n_T} \sim N\left(\frac{1}{n_T}\sum_{i=1}^{k} n_i\mu_i, \frac{\sigma^2}{n_T}\right)$ $\bar{X}_{i\cdot} = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i} \sim N\left(\boldsymbol{\mu_i}, \frac{\sigma^2}{n_i}\right)$
표본분산	$\boldsymbol{s^2} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{\cdot\cdot})^2}{n_T - 1} = MST = \frac{SST}{n_T - 1}$	

$\mathbf{SST} = \boldsymbol{(n_T - 1)s^2} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{\cdot\cdot})^2 = \sum_{i=1}^{k} n_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i\cdot})^2 = \mathbf{SSA + SSE}$

$\because \text{SST} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - n_T(\bar{X}_{\cdot\cdot})^2, \quad \text{SSA} = \sum_{i=1}^{k} n_i(\bar{X}_{i\cdot})^2 - n_T(\bar{X}_{\cdot\cdot})^2, \quad \text{SSE} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - \sum_{i=1}^{k} n_i(\bar{X}_{i\cdot})^2$

$$\boldsymbol{Q = Q_1 + Q_2} \Leftrightarrow \boldsymbol{(n_T - 1)S^2} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i\cdot})^2 + \sum_{i=1}^{k} n_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$$

	SST $(= \boldsymbol{Q})$	SSE $(= \boldsymbol{Q_1})$	SSA $(= \boldsymbol{Q_2})$
정의	$\boldsymbol{(n_T - 1)s^2}$	$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i\cdot})^2$	$\sum_{i=1}^{k} n_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$
χ^2	$\boldsymbol{\frac{\text{SST}}{\sigma^2}}$	$\boldsymbol{\frac{\text{SSE}}{\sigma^2}}$	$\boldsymbol{\frac{\text{SSA}}{\sigma^2}}$
	$\boldsymbol{H_0}: \frac{\text{SST}}{\sigma^2} = \frac{(n_T - 1)s^2}{\sigma^2} \sim \chi^2(n_T - 1)$	$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i\cdot})^2$ $\frac{\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i\cdot})^2}{\sigma^2} = \frac{(n_i - 1)s_i^2}{\sigma^2} \sim \chi^2(n_i - 1)$ $\therefore \frac{\text{SSE}}{\sigma^2} = \sum_{i=1}^{k}\frac{(n_i - 1)s_i^2}{\sigma^2} \sim \chi^2(n_T - k)$	$\boldsymbol{H_0}: \frac{\text{SSA}}{\sigma^2} = \frac{\text{SST}}{\sigma^2} - \frac{\text{SSE}}{\sigma^2} \sim \chi^2(k - 1)$
자유도	$\boldsymbol{n_T - 1}$	$\boldsymbol{n_T - k}$	$\boldsymbol{k - 1}$
Mean Square	$\mathbf{MST} = \frac{\mathbf{SST}}{\boldsymbol{n_T - 1}}$	$\mathbf{MSE} = \frac{\mathbf{SSE}}{\boldsymbol{n_T - k}}$	$\mathbf{MSA} = \frac{\mathbf{SSA}}{\boldsymbol{k - 1}}$
제곱합	총 제곱합	열(처리)-내부	열(처리) 평균간

*핵심 질문: 요인 k개의 평균이 모두 동일 한가? $(H_0: \mu_1 = \cdots = \mu_k)$

$$H_0: L(\omega) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n_T}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \boldsymbol{\mu})^2\right] \quad \omega = \{(\mu_1, \mu_2, \cdots, \mu_k, \sigma^2): -\infty < \mu_1 = \mu_2 = \cdots = \mu_k = \mu < \infty,\ 0 < \sigma^2 < \infty\}$$

$$H_1: L(\Omega) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n_T}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \boldsymbol{\mu_i})^2\right] \quad \Omega = \{(\mu_1, \mu_2, \cdots, \mu_k, \sigma^2): -\infty < \mu_i < \infty,\ 0 < \sigma^2 < \infty\}$$

➔ 각각 (μ, σ^2), $(\mu_1, \cdots, \mu_i, \cdots, \mu_k, \sigma^2)$에 대해 편미분 후 MLE 구하면

$$\Lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \left[\frac{1/\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{\boldsymbol{x}}_{\cdot\cdot})^2}{1/\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{\boldsymbol{x}}_{i\cdot})^2}\right]^{\frac{n_T}{2}} = \left[\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{\boldsymbol{x}}_{i\cdot})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{\boldsymbol{x}}_{\cdot\cdot})^2}\right]^{\frac{n_T}{2}} = \left[\frac{\text{SSE}}{\text{SST}}\right]^{\frac{n_T}{2}} \quad \therefore \Lambda^{2/n_T} = \frac{\text{SSE}}{\text{SSE} + \text{SSA}} = \frac{1}{1 + \text{SSA/SSE}}$$

$$\alpha = P_{H_0}\left[\frac{1}{1 + \text{SSA/SSE}} \le \lambda_0^{2/n_T}\right] = P_{H_0}\left[\boldsymbol{F} = \frac{\mathbf{MSA}}{\mathbf{MSE}} = \frac{\text{SSA}/(k-1)}{\text{SSE}/(n_T - k)} \ge \boldsymbol{c} = \frac{n_T - k}{k - 1}(\lambda_o^{-2/n_T} - 1)\right]$$

<**모델의 재구성**> 처리에 의한 변동 (SSA) / 내부 오차에 의한 변동을 χ^2로 표현 (SSE) → F-통계량 구현 시

$\boldsymbol{F} = \frac{\mathbf{MSA}}{\mathbf{MSE}} = \frac{\text{SSA}/(k-1)}{\text{SSE}/(n_T - k)} \ge c$ 이면 H_0 기각 (≈ 처리에 의한 변동이 임계치를 넘음)

k=2일 때 **등분산** $\boldsymbol{T^2 = F}$

$$pf)\ F = \frac{\text{MSA}}{\text{MSE}} = (n_1 + n_2 - 2)\frac{\text{SSA}}{\text{SSE}} = \frac{n_1\bar{X}_1^2 + n_2\bar{X}_2^2 - (n_1 + n_2)\left(\frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2}\right)^2}{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \frac{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)(\bar{X}_1 - \bar{X}_2)^2}{S_p^2} = \left\{\frac{\bar{X}_1 - \bar{X}_2}{S_p/\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right\}^2$$

<**Tukey 다중비교 절차**> ➔ 개별 비교 시 100(1-α)% 동시 신뢰구간 (q: Studentized range 분포)

$$\mu_{i_1} - \mu_{i_2} \in \left(\bar{x}_{i_1\cdot} - \bar{x}_{i_2\cdot} - \frac{\sqrt{\text{MSE}}\, q_{\alpha,k,n_T-k}}{\sqrt{2}}\sqrt{\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}},\ \bar{x}_{i_1\cdot} - \bar{x}_{i_2\cdot} + \frac{\sqrt{\text{MSE}}\, q_{\alpha,k,n_T-k}}{\sqrt{2}}\sqrt{\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}}\right) \quad \text{for } i_1, i_2 \in \{1,2,3,4\}$$

8. 정규모형 추론: ANOVA

2-Way ANOVA

1-way ANOVA: b개 처리 별 → 동일 크기 (a) 샘플 간 평균 동일성

2-way ANOVA: 요인 A (n=1, ..., a) + 요인 B (n=1, ..., b)

① $X_{ij} \overset{iid}{\sim} N(\mu_{ij}, \sigma^2),\ n = ab$

② $\mu_{ij} = \mu + (\bar{\mu}_{i\cdot} - \mu) + (\bar{\mu}_{\cdot j} - \mu) = \mu + \boldsymbol{\alpha_i} + \boldsymbol{b_j} \quad (\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = 0)$

*Additive model $(\because \sum_{j=1}^{b}\sum_{i=1}^{a} \mu_{ij} = \sum_{i=1}^{a} b\bar{\mu}_{i\cdot} + \sum_{j=1}^{b} a\bar{\mu}_{\cdot j} - ab\mu = ab\mu)$

	$\boldsymbol{a}$개$(\boldsymbol{i})$ **처리** 개수						
$\boldsymbol{b}$개$(\boldsymbol{j})$ 처리 개수	X_{11}	X_{12}	$\cdots$	X_{1j}	$\cdots$	X_{1b}	$\bar{X}_{1\cdot}$
	X_{21}	X_{22}	$\cdots$	X_{2j}	$\cdots$	X_{2b}	$\bar{X}_{2\cdot}$
	$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\vdots$	$\vdots$
	X_{i1}	X_{i2}	$\cdots$	X_{ij}	$\cdots$	X_{ib}	$\boldsymbol{\bar{X}_{i\cdot}}$
	$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\vdots$	$\vdots$
	X_{a1}	X_{a2}	$\cdots$	X_{aj}	$\cdots$	X_{ab}	$\bar{X}_{a\cdot}$
	$\bar{X}_{\cdot 1}$	$\bar{X}_{\cdot 2}$	$\cdots$	$\boldsymbol{\bar{X}_{\cdot j}}$		$\bar{X}_{\cdot b}$	$\boldsymbol{\bar{X}_{\cdot\cdot}}$

$$\boldsymbol{Q = Q_1 + Q_2} \iff \boldsymbol{(ab-1)S^2} = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(X_{ij} - \bar{X}_{i\cdot}\right)^2 + b\sum_{i=1}^{a}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$$

$$\boldsymbol{Q = Q_3 + Q_4} \iff \boldsymbol{(ab-1)S^2} = \sum_{j=1}^{b}\sum_{i=1}^{a}\left(X_{ij} - \bar{X}_{\cdot j}\right)^2 + a\sum_{j=1}^{b}\left(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}\right)^2$$

	$\boldsymbol{Q}$	$\boldsymbol{Q_1}$	$\boldsymbol{Q_2}$	$\boldsymbol{Q_3}$	$\boldsymbol{Q_4}$	$\boldsymbol{Q_5}$
정의	$\boldsymbol{(ab-1)S^2}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\left(X_{ij} - \bar{X}_{i\cdot}\right)^2$	$b\sum_{i=1}^{a}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$	$\sum_{j=1}^{b}\sum_{i=1}^{a}\left(X_{ij} - \bar{X}_{\cdot j}\right)^2$	$a\sum_{j=1}^{b}\left(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}\right)^2$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\left(X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot}\right)^2$
$\boldsymbol{\chi^2}$ 자유도 $\boldsymbol{Q_k/\sigma^2}$	$\boldsymbol{ab-1}$	$\boldsymbol{a(b-1)}$	$\boldsymbol{a-1}$	$\boldsymbol{b(a-1)}$	$\boldsymbol{b-1}$	$\boldsymbol{(a-1)(b-1)}$
제곱합	**총 제곱합**	**열-내부**	**열 평균간**	**행-내부**	**행 평균간**	**나머지 (Q2, Q4 제외)**

$$\boldsymbol{Q = Q_2 + Q_4 + Q_5} \iff \boldsymbol{(ab-1)S^2} = b\sum_{i=1}^{a}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 + a\sum_{j=1}^{b}\left(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}\right)^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\left(X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot}\right)^2$$

*핵심 질문: 요인 a,b의 effect ① $H_{0A}: \alpha_1 = \cdots = \alpha_a = 0$ (행간 차이 X)

② $H_{0B}: \beta_1 = \cdots = \beta_b = 0$ (열간 차이 X)

*열간 차이 모델 (H_{0B}, H_{1B})

$$H_{0B}: \hat{\sigma}^2_{\omega} = \frac{Q_4 + Q_5}{ab} = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\left(X_{ij} - \bar{X}_{i\cdot}\right)^2}{ab}$$

$$H_{1B}: \hat{\sigma}^2_{\Omega} = \frac{Q_5}{ab}$$

$$\Lambda = \left(\hat{\sigma}^2_{\Omega}/\hat{\sigma}^2_{\omega}\right)^{\frac{ab}{2}} \iff F = \frac{Q_4/(b-1)}{Q_5/[(a-1)(b-1)]} \geq \boldsymbol{c}$$

$*$ 행간 차이 모델 $(H_{0A}, H_{1A}) \iff F = \dfrac{Q_2/(a-1)}{Q_5/[(a-1)(b-1)]} \geq \boldsymbol{c}$

2-way ANOVA 일반화: 칸당 c>1개 관측값. (위 모델은 칸 당 1개) ➔ 총 n=abc (3차원 구조)

① $X_{ijk} \overset{iid}{\sim} N(\boldsymbol{\mu}_{ij}, \sigma^2),\ n = abc$

② $\mu_{ij} = \mu + (\bar{\mu}_{i\cdot} - \mu) + (\bar{\mu}_{\cdot j} - \mu) + \left(\mu_{ij} - \bar{\mu}_{i\cdot} - \bar{\mu}_{\cdot j} + \mu\right) = \mu + \boldsymbol{\alpha_i} + \boldsymbol{\beta_j} + \boldsymbol{\gamma_{ij}} \quad \left(\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = 0,\ \sum_{i=1}^{a} \gamma_{ij} = \sum_{j=1}^{b} \gamma_{ij} = 0\right)$

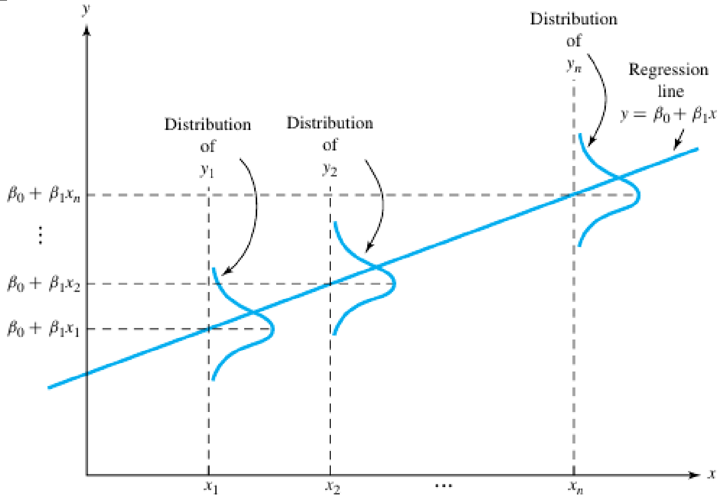
$\left(\gamma_{ij}: \text{교호작용 모수, interaction parameter}\right)$

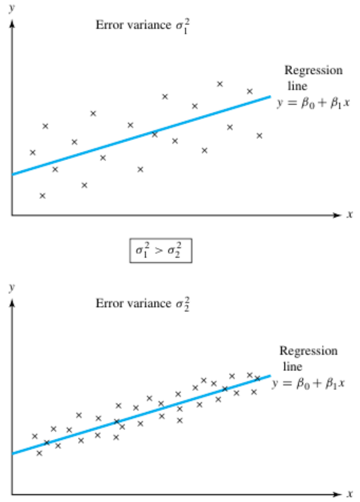
*핵심 가설: 교호작용 모수=0 $\ H_{0AB}: \boldsymbol{\gamma_{ij} = 0}$, for all $(i,j) \iff \boldsymbol{F}$통계량 $\geq \boldsymbol{c}$

9. 회귀분석

단순선형회귀

Fig





전제

1) $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ 🡐 Pairwise data: $(x_1, y_1), \cdots, (x_n, y_n)$

2) 오차: $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$ ⇔ $Y_i \overset{iid}{\sim} N(\beta_0 + \beta_1 x_i,\ \sigma^2)$ $\quad$ *$E(\varepsilon_i) = 0$ ⇔ $E(Y_i) = \beta_0 + \beta_1 x_i$

OLS

$$L(\beta_0, \beta_1, \sigma) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - (\beta_0 + \beta_1 x_i)}{\sigma}\right)^2\right\} = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\epsilon_i^2\right)$$

∴ 우도 최대 ⇔ $\sum_{i=1}^n \epsilon_i^2$ 최소 (OLS) ⇔ $\partial \sum \epsilon_i^2 / \partial \beta_0 = \partial \sum \epsilon_i^2 / \partial \beta_1 = \partial \sum \epsilon_i^2 / \partial \sigma = 0$ ⇔ MLE $(\hat\beta_0, \hat\beta_1, s^2_{MLE})$

추정치

$$* S_{XX} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2,\quad S_{YY} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2,\quad S_{XY} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$$

1) 회귀계수 추정: $\boldsymbol{\hat\beta_1 = \frac{S_{XY}}{S_{XX}},\quad \hat\beta_0 = \bar{y} - \hat\beta_1 \bar{x}}$

2) 관계식: ① $\bar{y} = \hat\beta_0 + \hat\beta_1 \bar{x}$ $\quad\leftrightarrow\quad$ $E(\bar{Y}) = \beta_0 + \beta_1 \bar{x}$ $\quad$ *$\bar{Y} \sim N\left(\beta_0 + \beta_1 \bar{x},\ \frac{\sigma^2}{n}\right)$

② $\hat{y}_i = \hat\beta_0 + \hat\beta_1 x_i$ $\quad\leftrightarrow\quad$ $E(Y_i) = \beta_0 + \beta_1 x_i$ $\quad$ $* \hat{y}_i \sim N\left(\beta_0 + \beta_1 x_i, \sigma^2\left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{XX}}\right)\right)$

③ $\hat{y}|x^* = \hat\beta_0 + \hat\beta_1 x^*$ $\quad\leftrightarrow\quad$ $E(Y|x^*) = \beta_0 + \beta_1 x^*$ $\quad$ $* \hat{y}|x^* \sim N\left(\beta_0 + \beta_1 x^*, \sigma^2\left(\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{XX}}\right)\right)$

3) $\text{SSE} = \sum_{i=1}^{n}\{y_i - (\hat\beta_0 + \hat\beta_1 x_i)\}^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}\hat{e}_i^2$ ($\hat{e}_i$: 잔차) $\left(\sum_{i=1}^{n}\epsilon_i^2\right.$ 에 MLE 값 대입$\left.\right)$

$s^2 = \frac{\text{SSE}}{n-2} = \text{MSE}$ (불편 추정량) $\quad cf)\ s^2_{MLE} = \frac{SSE}{n} = \frac{n-2}{n}s^2$

β_1

1) 정규성: $\boldsymbol{\hat\beta_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{XX}}\right)}$

2)검정/신뢰구간: $T_1 = \frac{\hat\beta_1 - \beta_1}{s/\sqrt{S_{XX}}} \sim \boldsymbol{t(n-2)}$ $\quad (H_0: \beta_1 = 0)$

$$\text{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(\epsilon_i - \bar\epsilon)^2 - \left(\hat\beta_1 - \beta_1\right)^2 S_{XX} \Leftrightarrow \sum_{i=1}^{n}\left(\frac{\epsilon_i}{\sigma}\right)^2 = \frac{\text{SSE}}{\sigma^2} + \left(\frac{\bar\epsilon}{\sigma/\sqrt{n}}\right)^2 + \left(\frac{\hat\beta_1 - \beta_1}{\sigma/\sqrt{S_{XX}}}\right)^2$$

$\therefore \frac{(n-2)s^2}{\sigma^2} = \frac{\text{SSE}}{\sigma^2} \sim \chi^2(n-2)$ $\quad \because \frac{\epsilon_i}{\sigma} \sim N(0,1),\ \bar\epsilon \sim N\left(0, \frac{\sigma^2}{n}\right),\ \hat\beta_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{XX}}\right)$

$$\hat\beta_1 = \frac{S_{XY}}{S_{XX}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{S_{XX}} = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)y_i$$

$$E(\hat\beta_1) = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)E(Y_i) = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)(\beta_0 + \beta_1 x_i) = \beta_0\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right) + \beta_1\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)x_i = 0 + \beta_1\frac{\sum x_i^2 - n\bar{x}^2}{S_{XX}} = \beta_1$$

$$\text{Var}(\hat\beta_1) = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)^2 \text{Var}(Y_i)^2 = \sigma^2\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)^2 = \frac{\sigma^2}{S_{XX}^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{\sigma^2}{S_{XX}}$$

β_0

1) 정규성: $\hat\beta_0 \sim N\left(\beta_0, \frac{\sum_{i=1}^n x_i^2}{nS_{XX}}\sigma^2\right)$

2)검정/신뢰구간: $T_0 = \frac{\hat\beta_0 - \beta_0}{s\sqrt{\sum_{i=1}^n x_i^2 / nS_{XX}}} \sim t(n-2)$

$$\text{Cov}(\hat\beta_0, \hat\beta_1) = Cov(\bar{y} - \bar{x}\hat\beta_1, \hat\beta_1) = Cov(\bar{y}, \hat\beta_1) - \bar{x}Var(\hat\beta_1) = Cov\left(\frac{1}{n}\sum_{i=1}^{n}y_i, \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)y_i\right) - \frac{\bar{x}\sigma^2}{Sxx}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{XX}}\right)Var(y_i) - \frac{\bar{x}\sigma^2}{Sxx} = -\frac{\bar{x}\sigma^2}{Sxx} \quad (\Leftrightarrow Cov(\bar{y}, \hat\beta_1) = 0)$$

9. 회귀분석

단순 선형 회귀

종속변수 기대값 추론 $E(Y|x^*)$

*모집단 $\boldsymbol{E(Y|x^*) = \beta_0 + \beta_1 x^*}$ ⇔ 점추정치 $\boldsymbol{\hat{y}|x^*} = \hat{\beta}_0 + \hat{\beta}_1 x^*$

$$\hat{y}|x^* \sim N\left(E(Y|x^*),\ \sigma^2\left(\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{XX}}\right)\right)$$

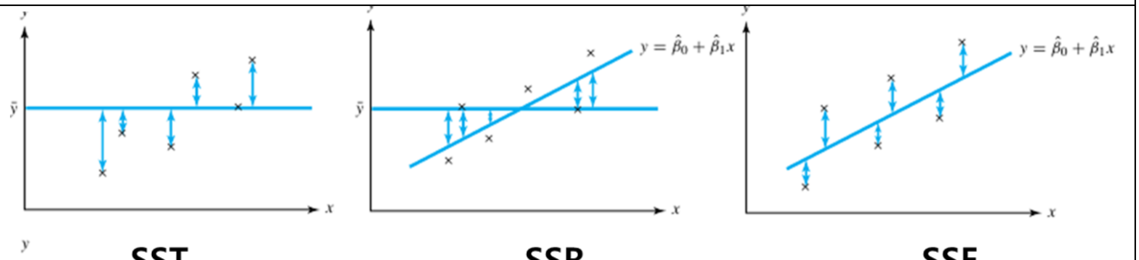
➔ 자유도 n-2인 t분포로 $E(Y|x^*)$의 신뢰구간 추론 가능

미래 예측 구간 $Y|x^*$

∗ 모집단 $\boldsymbol{Y|x^* = \beta_0 + \beta_1 x^* + \epsilon}$ ⇔ 점추정치 $\boldsymbol{\hat{y}|x^*} + \epsilon^* = \hat{\beta}_0 + \hat{\beta}_1 x^* + \epsilon^*$

$$\hat{\boldsymbol{y}}|\boldsymbol{x}^* + \epsilon^* \sim N\left(Y|x^*,\ \sigma^2\left(1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{XX}}\right)\right)$$

ANOVA



$$\underset{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\textbf{SST}} = \underset{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\textbf{SSR}} + \underset{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\textbf{SSE}}$$

변동 소스	d.f.	SS	MS	F-statistic
Regression	1	$\text{SSR}=\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \hat{\beta}_1^2 S_{XX} = \hat{\beta}_1 S_{XY}$	$\text{MSR} = \text{SSR}$	$F = \text{MSR}/\text{MSE}$
Error	$n-2$	$\text{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$\text{MSE} = \frac{\text{SSE}}{n-2}$	
Total	$n-1$	$\text{SST} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = S_{YY}$	$\text{MST} = \frac{\text{SST}}{n-1}$	

① $\frac{\text{SSE}}{\sigma^2} \sim \chi^2(n-2)$

② $\frac{\text{SSR}}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\{(\hat{\beta}_0 + \hat{\beta}_1 x_i) - (\hat{\beta}_0 + \hat{\beta}_1 \bar{x})\}^2 = \frac{\hat{\beta}_1^2}{\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \left(\frac{\hat{\beta}_1}{\sigma/\sqrt{S_{XX}}}\right)^2 \sim \chi^2(1)$

* $H_0: \beta_1 = 0$ 하에서 ... 만약 아니라면 $\frac{\text{SSR}}{\sigma^2} \sim \chi^2(1, \beta_1^2)$

$\boldsymbol{F = T_1^2}$

pf) $F = \frac{\text{MSR}}{\text{MSE}} = (n-2)\frac{\text{SSR}}{\text{SSE}} = S_{XX}\frac{\hat{\beta}_1^2}{s^2} = \left(\frac{\hat{\beta}_1}{s/\sqrt{S_{XX}}}\right)^2 = T_1^2$

결정계수

$$\boldsymbol{R^2} = \frac{\boldsymbol{SSR}}{\boldsymbol{SST}} = \hat{\beta}_1^2\frac{S_{XX}}{S_{YY}} = 1 - \frac{SSE}{SST} = \frac{1}{1 + \frac{SSE}{SSR}}$$

상관분석

$$\textbf{Pearson } \boldsymbol{r} = \frac{\boldsymbol{S_{XY}}}{\sqrt{\boldsymbol{S_{XX}}}\sqrt{\boldsymbol{S_{YY}}}} = \hat{\boldsymbol{\beta}}_1\frac{\sqrt{\boldsymbol{S_{XX}}}}{\sqrt{\boldsymbol{S_{YY}}}} \quad \left(\rho = \beta_1\frac{\sigma_X}{\sigma_Y},\ \begin{bmatrix} X_i \\ Y_i \end{bmatrix} \sim BVN\left(\begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}\right)\right)$$

$$T_1 = \frac{\hat{\beta}_1}{s/\sqrt{S_{XX}}} = \frac{r\sqrt{S_{YY}}}{s} = \frac{r\sqrt{S_{YY}}}{\sqrt{\frac{S_{YY} - \hat{\beta}_1^2 S_{XX}}{n-2}}} = \frac{r\sqrt{n-2}}{\sqrt{1 - \hat{\beta}_1^2\frac{S_{XX}}{S_{YY}}}} = \frac{\boldsymbol{r\sqrt{n-2}}}{\sqrt{\boldsymbol{1-r^2}}} \sim t(n-2) \quad (H_0: \rho = 0)$$

$$W = \frac{1}{2}\ln\left(\frac{1+R}{1-R}\right) \xrightarrow{D} N\left(\frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right), \frac{1}{n-3}\right)$$ ∗ 증명 없이 ...

잔차분석

$\hat{e}_i = y_i - \hat{y}_i \ \ (\sum_{i=1}^{n} e_i = 0)$

9. 회귀분석

다중선형회귀

전제

1) $y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \epsilon_i = \mathbf{x_i}^T\boldsymbol{\beta} + \epsilon_i$ ← Pairwise data: $(y_1, x_{11}, \cdots, x_{k1}), \cdots, (y_n, x_{1n}, \cdots, x_{kn})$

2) $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ ① $\mathbf{y} = [y_1, y_2, \cdots, y_n]^T$ ③ $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2, \cdots, \beta_k]^T$

② $\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} = \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \vdots \\ \mathbf{x_n} \end{bmatrix}$ ④ $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \cdots, \epsilon_n]^T$

3) 분포: $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I_n}) \Leftrightarrow \mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I_n})$

OLS

$$L(\boldsymbol{\beta}, \sigma) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - (\beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki})}{\sigma}\right)^2\right\} = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\epsilon_i^2\right)$$

∴ 우도 최대 ⇔ $\sum_{i=1}^{n}\epsilon_i^2 = \boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ 최소 ⇔ $\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{y}$ 해 ⇔ $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$

⇔ $\partial \sum \epsilon_i^2 / \partial\beta_0 = \cdots = \partial \sum \epsilon_i^2 / \partial\beta_k = \partial \sum \epsilon_i^2 / \partial\sigma = 0$ ⇔ MLE $(\widehat{\boldsymbol{\beta}}, s_{MLE}^2)$ (미분방정식도 같은 결과)

추정치

1) 관계식: ① $\bar{y} = \bar{\mathbf{x}}^T\widehat{\boldsymbol{\beta}}$ $\left(\bar{x} = \left[1, \frac{1}{n}\sum_{i=1}^{n} x_{1i}, \cdots, \frac{1}{n}\sum_{i=1}^{n} x_{ki}\right]^T\right)$ ↔ $\bar{y} = \bar{\mathbf{x}}^T\boldsymbol{\beta} + \bar{\epsilon}$

② $[\hat{y}_1, \cdots, \hat{y}_n]^T = \hat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$ $(\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \cdots + \hat{\beta}_k x_{ki} + \epsilon_i)$ ↔ $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$

$\hat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} = \{\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\}\mathbf{y} = \mathbf{H}\mathbf{y}$

*H: Hat matrix, 대칭 $(\mathbf{H} = \mathbf{H}^T)$ + 멱등원성 $(\mathbf{H}^k = \mathbf{H})$ + $(\mathbf{HX} = \mathbf{X})$; Col(X)에 y 정사영)

$\hat{y}_i = (\mathbf{x}_i)^T\widehat{\boldsymbol{\beta}}$ $(\mathbf{x}_i = [1, x_{1i}, \cdots, x_{ki}]^T)$ ↔ $E(\hat{y}_i) = (\mathbf{x}_i)^T\boldsymbol{\beta}$

③ $\hat{y}|\mathbf{x}^* = (\mathbf{x}^*)^T\widehat{\boldsymbol{\beta}}$ $(\mathbf{x}^* = [1, x_1^*, \cdots, x_k^*]^T)$ ↔ $E(Y|\mathbf{x}^*) = (\mathbf{x}^*)^T\boldsymbol{\beta}$

2) $\mathrm{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2 = \mathbf{e}^T\mathbf{e} = \mathbf{y}^T(\mathbf{I}_n - \mathbf{H})\mathbf{y}^T$ ($\mathbf{e} = [e_1, \cdots, e_n]^T = (\mathbf{I}_n - \mathbf{H})\mathbf{y}$: 잔차 벡터)

$s^2 = \frac{\mathrm{SSE}}{n-k-1} = \mathrm{MSE}$ (불편 추정량) $cf)$ $s_{MLE}^2 = \frac{SSE}{n} = \frac{n-k-1}{n}s^2$

주요 분포

① $\widehat{\boldsymbol{\beta}} \sim N_{k+1}(\boldsymbol{\beta}, \boldsymbol{\sigma}^2(\mathbf{X}^T\mathbf{X})^{-1})$ ∵ $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \sim N_{k+1}((\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}, (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\sigma^2\mathbf{I}_n)\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}) = N_{k+1}(\boldsymbol{\beta}, \sigma^2(\mathbf{X}^T\mathbf{X})^{-1})$

$\frac{\hat{\beta}_i - \beta_i}{\mathrm{se}(\hat{\beta}_i)} \sim t(n-p-1)$ $(\boldsymbol{H_0}: \boldsymbol{\beta_i} = \mathbf{0})$ * $\mathrm{se}(\hat{\beta}_i)$는 $(\mathbf{X}^T\mathbf{X})^{-1}$ i번째 대각 성분에 s 곱함

② $\hat{\mathbf{y}} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{H})$ ∵ $\mathbf{Hy} \sim N_n(\mathbf{HX}\boldsymbol{\beta}, \mathbf{H}(\sigma^2\mathbf{I}_n)\mathbf{H}^T) = N_n(\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{H}^2) = N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{H})$

- $\hat{y}_i \sim N(\mathbf{x}_i^T\boldsymbol{\beta},\ \sigma^2\{\mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})\mathbf{x}_i\})$
- $\hat{y}|\mathbf{x}^* \sim N(((\mathbf{x}^*)^T\boldsymbol{\beta},\ \sigma^2\{(\mathbf{x}^*)^T(\mathbf{X}^T\mathbf{X})\mathbf{x}^*\})$ ← 반응변수 기대값
- $\hat{y}|\mathbf{x}^* + \epsilon^* \sim N(((\mathbf{x}^*)^T\boldsymbol{\beta},\ \sigma^2\{1 + (\mathbf{x}^*)^T(\mathbf{X}^T\mathbf{X})\mathbf{x}^*\})$ ← 반응변수 예측

③ $\hat{\mathbf{e}} \sim N_n(\mathbf{0}, \sigma^2(\mathbf{I}_n - \mathbf{H}))$ ∵ $(\boldsymbol{I}_n - \mathbf{H})\boldsymbol{y} \sim N_n((\boldsymbol{I}_n - \mathbf{H})\mathbf{X}\boldsymbol{\beta}, (\boldsymbol{I}_n - \mathbf{H})(\sigma^2\mathbf{I}_n)(\boldsymbol{I}_n - \mathbf{H})^T) = N_n(\mathbf{0}, \sigma^2(\mathbf{I}_n - \mathbf{H}))$

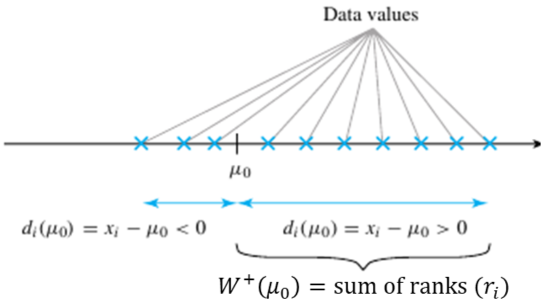
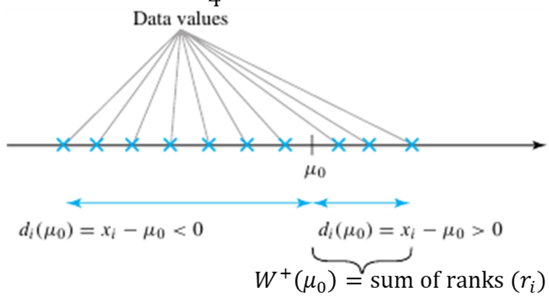
ANOVA

변동 소스	d.f.	SS	MS	F-statistic
Regression	k	$\mathrm{SSR} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	$\mathrm{MSR} = \frac{\mathrm{SSR}}{k}$	$F = \mathrm{MSR}/\mathrm{MSE}$
Error	$n-k-1$	$\mathrm{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$\mathrm{MSE} = \frac{\mathrm{SSE}}{n-k-1}$	
Total	$n-1$	$\mathrm{SST} = \sum_{i=1}^{n}(y_i - \bar{y})^2$	$\mathrm{MST} = \frac{\mathrm{SST}}{n-1}$	

① $\frac{\mathrm{SSE}}{\sigma^2} \sim \chi^2(n-k-1)$

② $\frac{\mathrm{SSR}}{\sigma^2} = \sim \chi^2(k)$ * $\boldsymbol{H_0}: \boldsymbol{\beta_1} = \cdots = \boldsymbol{\beta_k} = \mathbf{0}$

10. 비모수통계

기본	1) Note: ①분포 무관 방법 (distribution-free) ② 모집단 가정: 연속성, 대칭성 ③ 2 개 이상 분포 함수의 집합에서도 가능함 2) 통계적 추론 Review ① 추정: 모집단 모수 추정 (점 추정, CI 추정) ② 가설 검정: H_0, H_1 설정 → 검정통계량 분포 (under H_0) → α(유의 수준)에서 p-value (유의 확률) 이용 기각 3) 점근상대효율 (Asymptotic Relative Efficiency; ARE): 의 효율성 ① **추정량/CI**: $ARE(\hat{\theta}_1,\hat{\theta}_2)=\lim_{n\to\infty}\frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}=\lim_{n\to\infty}\frac{l_2^2}{l_1^2}=ARE(CI_1,CI_2)$ $\quad * \, l_2=U_2-L_2$ when $CI_2\in(L_2,U_2)$ ② **검정**: $ARE(T_1,T_2)=\lim_{\theta_1\to\theta_0}\frac{n_2}{n_1}$ (required sample sizes under H_1) ⇒ **표본이 특정분포 따를 때, ARE로 효율성 결정**
분포	범함수를 활용하여 ECDF (추정)인 $\hat{F}(x_0)$ → Unknown CDF $F(x_0)$를 점추정 ① $\hat{F}(x_0)=\frac{\#(X_i\le x_0)}{n}$ (계단형 증가함수) $\xrightarrow{P} F(x_0)$ ② $S(x_0)=\#(X_i\le x_0)=\sum_{i=1}^{n} I(X_i\le x_0)\sim B(n,F(x_0))$ ⇔ $E\left(\hat{F}(x_0)\right)=\frac{E(S(x_0))}{n}=F(X_0)$, $\mathrm{Var}\left(\hat{F}(x_0)\right)=\frac{F(x_0)(1-F(x_0))}{n}$
1 표본 **Sign test**	*위치 모수 모형: $X_i=\theta+\epsilon_i$ $\left(\epsilon_i\overset{iid}{\sim}F,\ \text{which has location parameter } 0\right)$ 1) 가설: $H_0:\theta=\theta_0\ \ vs. H_1:\theta\neq\theta_0$ (>,< 가능) ① 가장 간단한 위치모수 비모수검정법 ② 이항분포 이용 2) 검정통계량: $S(\theta_0)$ ① 이항 정확성: $2\min\{P_{H_0}(X\ge S(\theta_0)),P_{H_0}(X\le S(\theta_0))\}$ ② 정규근사: Median 의 경우 $S(\theta_0)\sim B(n,F(\theta_0))=B\left(n,\frac{1}{2}\right)\approx N\left(\frac{n}{2},\frac{n}{4}\right)$
1 표본 **Wilcoxon Signed Rank**	* $(X_i-\theta_0)$의 부호+상대적인 크기 모두 고려한 검정법 1) 가정: **대칭성**, ∴ 평균 = 중위수, $E(X)=\theta_0=F^{-1}(1/2)$ 2) 가설: $H_0:\mu=\mu_0\ \ vs. H_1:\mu\neq\mu_0$ (>,< 가능) 3) 전개: $d_i=X_i-\mu_0$에 대해 $\lvert d_i\rvert$의 순위 등수 점수 r_i ($if\ \lvert d_i\rvert$가 3 등 ⇒ $r_i=3$) *공동 6 위 2 명 → 6.5 위 ($\sum_{i=1}^{n} r_i=n(n+1)/2$, 만약 $d_i=0$ ⇔ $X_i=\mu_0$인 자료 있으면 삭제 후 n′ = n − 1) 4) 검정통계량: $W^+=\sum_{i=1}^{n}\{I(d_i)\}\boldsymbol{r_i}=\sum_{d_i>0}\boldsymbol{r_i}$ → $W^+>n(n+1)/4$ 이면 $\mu>\mu_0$ $W^+<n(n+1)/4$ 이면 $\mu<\mu_0$ 5) 정규 근사: $W^+=\sum_{k=1}^{n} kW_k\xrightarrow{D}N\left(\frac{n(n+1)}{4},\frac{n(n+1)(2n+1)}{24}\right)$, 연속성 수정 가능 (동점자 없을 때… k 는 각 rank, W_k는 포함(1) or 불포함(0)의 베르누이) ∵ $W_k\sim B\left(1,\frac{1}{2}\right)$ ⇒ ① $E(W^+)=E\left(\sum_{k=1}^{n}kW_k\right)=\sum_{k=1}^{n}kE(W_k)=\sum_{k=1}^{n}\frac{k}{2}=\frac{n(n+1)}{4}$ ② $\mathrm{Var}(W^+)=\sum_{k=1}^{n}k^2\mathrm{Var}(W_k)=\sum_{k=1}^{n}\frac{k^2}{4}=\frac{n(n+1)(2n+1)}{24}$ 6) 정확성 검정: $W^+=\sum_{k=1}^{n}kW_k=x$ 가 되는 $\{W_k\}$ 조합의 수가 m 개 ⇔ $P(W^+=x)=m/2^n$
	$W^+(\mu_0)>\frac{n(n+1)}{4}\ \Rightarrow\ \mu>\mu_0$  $W^+(\mu_0)<\frac{n(n+1)}{4}\ \Rightarrow\ \mu<\mu_0$ 

10. 비모수통계

2 표본 **Mann-Whitney-Wilcoxon (MWW) U** **=** **Wilcoxon Rank-sum**	*X,Y가 같은 함수 but 다른 위치모수를 가질 때 ...$X_i=\theta+\epsilon_i$, $Y_i=\theta+\Delta+\epsilon_i$ 1) 가정: $X_i \overset{iid}{\sim} \boldsymbol{F}(\theta)$, $Y_j \overset{iid}{\sim} \boldsymbol{F}(\theta+\Delta)$ 2) 가설: $H_0:\Delta=0$, $H_1:\Delta\neq 0$ (>,< 가능) 3) 전개: n개의 $\{X_i\}$와 m개의 $\{Y_j\}$ → $(n+m)$개의 순위 값을 매긴다. 4) 검정통계량: $W_X=\sum_i \boldsymbol{r_i}$ (혼합 표본에서 X_i들의 순위 r_i합) ⇒ W_X의 범위: $\left(\frac{n(n+1)}{2},\left\{mn+\frac{n(n+1)}{2}\right\}\right)$ ① $E(W_X)=\frac{n}{2}(m+n+1)$ ② $\mathrm{Var}(W_X)=\frac{mn}{12}(m+n+1)$ 5) 정규근사: $W_X \xrightarrow{D} N\left(\frac{n}{2}(m+n+1), \frac{mn}{12}(m+n+1)\right)$ 6) U통계량: $U=W_X-\frac{n(n+1)}{2}\xrightarrow{D}N\left(\frac{mn}{2}, \frac{mn}{12}(m+n+1)\right)$ ⇒ U의 범위: $(0,mn)$ 7) 정확성 검정: Specific order of X in pooled sample ex) $[1,1,0,0,1,\cdots,1]$ 위 순서가 발생할 확률 $P=\frac{1}{\binom{n+m}{n}}$ ⇒ W_X or U 값에 대해 이산 분포 그려서 P값 산출
	$*\ \mathrm{n+m=N}$ 일 때 귀무가설 하: $E(r_i)=\frac{1}{N}\sum_{k=1}^{N}k=\frac{N+1}{2}$, $E(r_i^2)=\frac{1}{N}\sum_{k=1}^{N}k^2=\frac{(N+1)(2N+1)}{6}$ $\sum_{j=1}^{N}\sum_{k=1}^{N}jk=\left(\sum_{k=1}^{N}k\right)^2=\frac{N^2(N+1)^2}{4}$ ⇒ $\sum_{j\neq k}jk=\frac{N^2(N+1)^2}{4}-\frac{N(N+1)(2N+1)}{6}$ $\Rightarrow E[r_ir_j]_{j\neq k}=\frac{1}{N(N-1)}\sum_{j\neq k}jk=\frac{N(N+1)^2}{4(N-1)}-\frac{(N+1)(2N+1)}{6(N-1)}$ $Var(r_i)=E(r_i^2)-E(r_i)^2=\frac{N^2-1}{12}$, $Cov(r_i,r_j)=E[r_ir_j]-E(r_i)E(r_j)=-\frac{N+1}{12}$ $\therefore E(W_X)=E\left(\sum_{i=1}^{n}r_i\right)=\frac{n(N+1)}{2}=\frac{n(m+n+1)}{2}$ $\therefore Var(W_X)=\sum_{i=1}^{n}Var(r_i)+2\sum_{i<j}Cov(r_i,r_j)=n\left(\frac{N^2-1}{12}\right)-n(n-1)\left(\frac{N+1}{12}\right)=\frac{mn}{12}(m+n+1)$
2 표본 **Kolmogorov-Smirnov**	2 개 분포의 동등성에 대한 검정 or 1 개 분포의 적합도 검정 1) 전개 ① $X_i\sim F$, $Y_j\sim G$ ② $\hat F(x)=\frac{\#(x_i\le x)}{n}$, $\hat G(y)=\frac{\#(y_i\le y)}{m}$ 2) 가설: $H_0:F(x)=G(x)$ *2 개 분포 (CDF)는 동등하다. 3) 검정통계량: $M=\max\left\|\hat F(x)-\hat G(x)\right\|$ (두 ECDF 의 최대 수직거리) ⇒ Reject H_0 if $M>d_\alpha\sqrt{\frac{1}{n}+\frac{1}{m}}$
≥3 표본 **Kruskal-Wallis test**	3 개 이상 treatment 의 **위치모수** 동일성 검정 1) 가정 $x_{ij}=\mu_j+\epsilon_{ij}$, $\epsilon_{ij}\overset{iid}{\sim}F$ 2) 가설: $H_0:\mu_1=\cdots=\mu_k$ 3) 전개: k개 treatment 전부 합쳐서 n_T개 sample 만듦 → 순위 매겨서 순위합 $\frac{n_T(n_T+1)}{2}$ 표본평균 순위합 산출 가능 *$\bar r_{i.}=(r_{i1}+\cdots+r_{in_i})/n_i$ 4) 검정통계량: $H=\frac{12}{n_T(n_T+1)}\sum_{i=1}^{k}n_i\bar r_{i.}^{\,2}-3(n_T+1)\sim\boldsymbol{\chi^2}(k-1)$ 5) 사후 검정: Dunnett test
연관성 척도	① **Spearman's rank correlation (ρ)**: 임의의 이변량 분포, 피어슨 표본상관계수를 **순위 (rank)**로 재계산 $r_S=\dfrac{\sum_{i=1}^{n}\left\{R(X_i)-\frac{n+1}{2}\right\}\left\{R(Y_i)-\frac{n+1}{2}\right\}}{\frac{n(n^2-1)}{12}}$ ② **Kendall's tau (τ)**: 임의의 이변량 분포. $(X_i-X_j)(Y_i-Y_j)$ 부호로 concordance 정함... (**두 변수가 단조적인가**?) $\tau=P(concordance)-P(discordance)=2P_c-1,\therefore\ -1\le\tau\le1$ $K=\binom{n}{2}^{-1}\sum_{i<j}\mathrm{sgn}\{(X_i-X_j)(Y_i-Y_j)\}$, where $E(K)=\tau$, $K\xrightarrow{D}N\left(0,\frac{2(2n+1)}{9n(n-1)}\right)$ under $H_0:\tau=0$

11. 베이지안 통계학 기본

베이지안 절차	**사전/사후 분포**	①기본: $X_i\|\theta \overset{iid}{\sim} f(x\|\theta)$ (X는 θ에 의존적인 확률 분포에서 추출) ②사전분포: $\Theta \sim h(\theta)$ *모수의 **prior** ③우도: $L(\mathbf{x}\|\theta) = f(x_1\|\theta)\cdots f(x_n\|\theta)$ $*$ 표본 $\mathbf{X}^T = (X_1,\cdots,X_n)$ ④결합 PDF: $g(\mathbf{x},\theta) = L(\mathbf{x}\|\theta)h(\theta)$ ∴ 사후분포: $k(\theta\|\mathbf{x}) = \frac{g(\mathbf{x},\theta)}{g_1(\mathbf{x})} = \frac{L(\mathbf{x}\|\theta)h(\theta)}{g_1(\mathbf{x})} = \frac{L(\mathbf{x}\|\theta)h(\theta)}{\int_{-\infty}^{\infty} L(\mathbf{x}\|\theta)h(\theta)\,d\theta}$ $\Rightarrow k(\theta\|\mathbf{x}) \propto L(\mathbf{x}\|\theta)h(\theta)$ *** 공액분포족: 사전⇔사후 분포가 같은 족에 속함** (e.g. X|θ가 푸아송 → 공액족 θ는 감마분포) X|θ가 이항분포 → 공액족 θ는 베타분포)
		<다변량 조건부/주변 분포 이용한 유도) 1) $h(\theta) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\cdots,x_n,\theta)\,dx_1\cdots dx_n$ 2) $L(\mathbf{x}\|\theta) = L_{\mathbf{x}\|\theta}(x_1,\cdots,x_n\,\|\theta) = g(\mathbf{x},\theta)/h(\theta)$ 3) $g_1(\mathbf{x}) = g_1(x_1,x_2,\cdots,x_n) = \int_{-\infty}^{\infty} g(\mathbf{x},\theta)\,d\theta$ (초등 적분으로 해결 안됨 ➔ **MCMC 사용**)
	베이지안 추정	1) 점추정: **Maximum A Posteriori** (**MAP**): 사후확률은 최대화하는 값으로 모수 추정 ($\hat{\theta}_{MAP}$) ① MLE는 모수 미지의 정해진 분포에서 iid 추출한 값들 → 샘플 기준으로 모수 추정 ∴ $\hat{\theta}_{MLE} = \underset{\theta}{\operatorname{argmax}} L(\mathbf{x}\|\theta)$ ② **MAP**는 모두~사전분포 + iid 추출 샘플들 → 두 가지를 혼합한 사후 분포에서 모수 추정 ∴ $\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}} L(\mathbf{x}\|\theta)h(\theta)$ ③ **균등 Prior [=Beta(1,1)]** → $h(\theta)$가 상수이므로 $\underset{\theta}{\operatorname{argmax}} L(\mathbf{x}\|\theta)h(\theta) = \underset{\theta}{\operatorname{argmax}} L(\mathbf{x}\|\theta)$ ⇔ $\boldsymbol{\hat{\theta}_{MLE} = \hat{\theta}_{MAP}}$ (기타 점추정 방식: 1) MAP, 2) 사후 평균, 3) 사후 중간값 등)
		2) 구간 추정: $P[u(\mathbf{x}) < \Theta < v(\mathbf{x})\ \|\ \mathbf{X} = \mathbf{x}] = \int_{u(\mathbf{x})}^{v(\mathbf{x})} k(\theta\|\mathbf{x})\,d\theta = 0.95$ ⇔ Credible interval (신용 구간) or 사후확률구간 0.95 ① HPDI: 사후분포에서 가장 짧은 길이 / ② Equal-tail CI (양측 꼬리 넓이 동일)
	베이지안 검정	$P(\Theta \in \omega_0\|\mathbf{x})\ vs.\ P(\Theta \in \omega_1\|\mathbf{x})$ ➔ **더 큰 쪽으로 가설 채택** (사후 분포 상 가설의 모수 영역 더 넓은 쪽)
	Hierarchical Bayes	① $X\|\theta \sim f(x\|\theta)$ ② $\Theta\|\gamma \sim h(\theta\|\gamma)$ ③ $\Gamma \sim \psi(\gamma)$ (γ: 초모수) $k(\theta\|\mathbf{x}) = \frac{\int_{-\infty}^{\infty} L(\mathbf{x}\|\theta)\,h(\theta\|\gamma)\,\psi(\gamma)\,d\gamma}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L(\mathbf{x}\|\theta)\,h(\theta\|\gamma)\,\psi(\gamma)\,d\gamma\,d\theta}$
	Empirical Bayes	① $X\|\theta \sim f(x\|\theta)$ ② $\Theta\|\gamma \sim h(\theta\|\gamma)$ $m(\mathbf{x}\|\gamma) = \int_{-\infty}^{\infty} g(\mathbf{x},\theta\|\gamma)\,d\theta = \int_{-\infty}^{\infty} L(\mathbf{x}\|\theta)\,h(\theta\|\gamma)\,d\theta$ ➔ 새로운 우도 최대화 하는 $\hat{\gamma}$ 구함 ➔ 위 조건부 사전분포에 대입하여 진행

*몬테카를로 기법/깁스샘플러