

수리통계 1. 2장 추가 문제

[추가 1] (조건부 기댓값의 성질)

조건부 기댓값의 성질을 이용하여 다음이 성립함을 밝혀라.

$$\text{Cov}(X, Y) = E[\text{Cov}(X, Y|Z)] + \text{Cov}[E(X|Z), E(Y|Z)]$$

[추가 2] (다차원 확률변수의 최소평균제곱선형예측자)

다차원 확률변수 X 의 1차 함수 $AX+b$ 로서

$$E[\|Y - (AX+b)\|^2]$$

을 최소로 하는 1차 함수는 $\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X - \mu_1)$ 이고, 그 최소값은

$$\text{trace}(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

임을 밝혀라. 여기에서 X 의 분산행렬은 역행렬을 갖는 정칙행렬임이 전제되어 있다.

[추가 3] (중상관계수(multiple correlation)의 뜻)

다차원 확률변수 X 의 실수값 1차 함수 a^tX+b 와 1차원 확률변수 Y 사이의 상관계수에 대하여 다음이 성립함을 밝히고, 그 최대값은 X 를 이용한 최소평균제곱선형예측자와 Y 사이의 상관계수임을 확인하여라.

$$\max_{a,b} \text{Corr}(Y, a^tX+b) = \left(\frac{\sigma_{21}\Sigma_{11}^{-1}\sigma_{12}}{\sigma_{22}} \right)^{1/2}$$

여기에서 X 의 분산행렬 Σ_{11} 의 역행렬 Σ_{11}^{-1} 이 존재하고, Y 의 분산 σ_{22} 는 양수임이 전제되어 있다.