기말고사 해답 및 채점 기준

담당교수: 전종준

2018년 12월 12일, 10:30 ~ 11:30

(I) 주어진 질문에 대해 답하여라 (각 10점, 부분점수 있음)

1. 확률밀도함수를 통해 알 수 있는 분포에 대한 정보를 히스토그램을 통해 설명하여라.

   (1) 확률밀도함수가 히스토그램의 극한이다. (3pt)

   (2) 특정부분에서의 히스그램의 막대의 상대적인 넓이가 데이터가 특정구간에 속할 상대 빈도다. (2pt)

   (3) 임의의 구간에 대한 상대빈도의 극한은 해당 확률이므로, 확률밀도함수를 통해 확률분포를 알 수 있다. (5pt)

2. 확률 분포를 설명하여라. (필요한 개념과 정의가 유기적으로 연결되어 있어야 함)

   - 실험, 표본공간, 확률, 확률변수, 실수위에서의 확률의 흩어짐. (각 2pt씩)

3. 어떤 확률변수의 누적분포함수의 정의를 쓰고 증가함수 임을 설명하여라.

   - 정의 (5pt)
   - 그림을 그리면 (2pt), 확률의 공리를 사용하면 (5pt)

4. 어떤 자료가 어떤 분포 F를 따른다는 말을 설명하여라.

   - 자료가 랜덤하다는 것이 표현되어 있나? (3pt)
   - 자료가 생성모형의 확률변수로 실현값임이 나타나 있나?(2pt)
   - 누적분포함수 혹은 확률밀도 함수로 자료의 분포를 설명할 수 있음을 기술하였는가? (5pt)

   (답안 예) 자료가 분포 F를 따른다는 것은 자료의 랜덤한 성질을 가지고 있다는 것을 가정한다. 어떤 확률변수 X가 있고 그것의 분포함수 (혹은 확률밀도 함수)를 F라고 할 때, 확률변수 X의 실현값을 관찰된 자료라고 가정하는 것이다.

5. 조건부 기대값의 의미를 실제 자료 분석의 예를 들어 설명하여라.

   - 조건부 기대값 정의를 올바로 명시하였나. 그리고 그 조건부 기대값이 predictor 의 함수로 표시되어 있나? $g(X) = E(Y|X)$ (5pt)
   - 자료분석과 연결하였나? (5pt)

6. 두 확률변수의 분포가 같다는 뜻을 설명하여라.

   - 두 확률변수의 분포를 표시하였는가? (5pt)
   - cdf, pdf 혹은 실수위 집합에 대해 확률값이 같음을 표시하였나? (5pt)

(II) 다음을 계산하여라. (각 20점, 부분점수 있음)

- 경험적 위험 최소화 방법을 다음 개념을 포함하여 설명하여라. (자료, 기대값, 손실함수, 위험함수, 대수의 법칙)

  (답안 예) 어떤 반응변수-설명변수에 대응되는 확률변수 쌍 (Y, X)가 주어져있다. 여기서 $\beta \in \mathbb{R}^p$에 의해 모수화 된 의사결정함수 $\delta(X;\beta)$가 주어져 있고 손실함수 $l(Y, \delta(X;\beta))$를 가정하자. 위험함수는 $E(l(Y, \delta(X;\beta)))$로 주어지며, 모집단에서의 추정량의 이 위험함수를 최소화 하는 β가 된다. 하지만 이 위험함수는 실제 계산할 수 없는 값으로 대응되는 확률변수 쌍 (Y, X)의 실현값(관측치)인 $(y_i, x_i)\quad, i = 1, \cdots, n$ 을 이용하여 경험적 위험함수

  $$\frac{1}{n}\sum_{i=1}^{n} l(y_i, \delta(x_i;\beta))$$

  를 최소화하는 β를 찾는다. 경험적 위험함수는 대수의 법칙에 의해 위험함수의 근사로 설명할 수 있고, 이는 경험적 위험함수가 실제 위험을 최소화 하는 β를 찾는 근사방법을서 이론적 근거가 된다.

- 선형회귀모형을 설명하고, 추정량을 구하는 방법을 다음 개념을 포함시켜 설명하여라. (선형 모형, 추정 대상, 행렬, 미분, 경험적 위험함수)

  (답안 예) 선형회귀모형은 $y = x^\top\beta + \epsilon$ 의 생성모형을 가정한다 (단, $\epsilon \perp x$, $\epsilon \sim (0, \sigma^2)$, y, x는 확률변수). 이 모형에서 조건부 기대값

  $$\mathrm{E}(y|x) = x^\top\beta$$

  이며 추정대상은 x가 주어진 경우 y의 조건부 기대값이된다. 이는 β를 통해 알 수 있으며, 제곱손실함수를 사용했을 때 유도되는 위험함수의 최소값이 참 β됨을 알 수 있다. 위험함수는

  $$R(\beta) = \mathrm{E}(y - \mathbf{x}^\top\boldsymbol{\beta})^2$$

  로 주어진다. 그러나 실제로 이 위험함수는 실제 계산이 불가능 한 값으로 경험적 위험함수로 대체하여 β의 추정값을 이용한다. Y를 y의 n개의 iid copy (혹은 관측치)인 n차원 벡터 X를 x의 iid copy (혹은 관측치)인 $n \times p$행렬이라고하면, $R(\beta)$는 경험적 위험함수는

  $$\hat{R}(\beta) = \frac{1}{n}\|Y - X\beta\|^2$$

  로 주어진다. 경험적 위험함수의 최소값은 미분을 통해 $\hat{\beta} = (X^\top X)^{-1}X^\top Y$임을 알 수 있다.

문제 끝.