

□ 통계 용어 점검 문제 50제

① 확률분포·확률변수

1. 확률변수란 무엇인가?

→ 결과에 수치를 부여한 변수 (불확실한 사건을 수로 표현)

2. 이산형 확률분포의 대표 예시 2가지는?

→ 베르누이 분포, 이항분포

3. 연속형 확률분포의 대표 예시 2가지는?

→ 정규분포, 지수분포

4. 이항분포의 모수(parameter)는?

→ 시행 횟수 n , 성공 확률 p

5. 포아송 분포는 어떤 상황에 쓰이는가?

→ 단위시간(또는 구간) 내 발생 횟수 모델링

6. 지수분포는 어떤 상황에 쓰이는가?

→ 사건 간 시간(대기시간) 모델링

7. 정규분포의 특징 세 가지는?

→ 대칭, 평균=중앙값=최빈값, 종 모양

8. 중심극한정리의 의미는?

→ 표본평균은 표본 크기가 커지면 정규분포를 따른다

9. 자유도가 중요한 분포 3가지는?

→ t 분포, χ^2 분포, F 분포

10. χ^2 분포는 주로 어디에 쓰이는가?

→ 분산 검정, 독립성 검정, 적합도 검정

② 기술통계·추정

11. 표본평균의 표준오차(Standard Error)는 무엇을 의미?
→ 표본평균의 변동성 (모평균 추정의 정확도)
12. 표준편차와 표준오차의 차이는?
→ 표준편차: 데이터의 흩어짐 / 표준오차: 평균 추정의 불확실성
13. 모평균의 점추정치로 사용하는 값은?
→ 표본평균
14. 모분산의 점추정치로 사용하는 값은?
→ 표본분산
15. 구간추정에서 신뢰구간의 의미는?
→ 모수가 포함될 가능성이 있는 구간
16. 신뢰수준 95%란 무엇을 의미?
→ 같은 방법을 반복했을 때 95% 구간이 모수를 포함
17. 모비율 추정에서 표준오차 공식은?
→ $\sqrt{p(1-p)/n}$
18. 큰 표본에서 모평균 추정에 사용하는 분포는?
→ 정규분포 (Z분포)
19. 모분산을 모르는 소표본에서 모평균 추정 시 사용하는 분포는?
→ t분포
20. 추정량이 불편추정량(unbiased estimator)이라는 의미는?
→ 평균적으로 모수를 정확히 추정한다

③ 가설검정

21.가설검정에서 H_0 , H_1 의미는?

→ H_0 : 귀무가설, H_1 : 대립가설

22.1종 오류(Type I error)란?

→ 사실은 H_0 가 맞는데 기각하는 오류

23.2종 오류(Type II error)란?

→ 사실은 H_1 이 맞는데 기각하지 못하는 오류

24.유의수준 α 란 무엇을 의미?

→ 1종 오류를 허용하는 확률

25.p-value가 유의수준보다 작으면?

→ 귀무가설 기각

26.두 집단 평균 비교 (모분산 σ^2 알려진 경우)는 어떤 검정 사용?

→ Z검정

27.두 집단 평균 비교 (모분산 σ^2 모르는 경우, 소표본)는 어떤 검정?

→ t검정

28.세 집단 이상 평균 비교 시 사용하는 방법은?

→ 분산분석 (ANOVA)

29.두 집단 분산 비교에 사용하는 검정은?

→ F검정

30. χ^2 독립성 검정은 어떤 상황에서 쓰이는가?

→ 범주형 변수 간 독립성 여부 판단

④ 상관·회귀분석

31.상관계수 r 의 범위는?

→ $-1 \leq r \leq 1$

32.상관계수 $r=0$ 의 의미는?

→ 선형 관계 없음

33.상관계수 $r=0.9$ 의 의미는?

→ 강한 양의 상관

34.상관은 인과관계와 같은가?

→ 아니다 (상관 \neq 인과)

35.단순회귀분석에서 종속변수와 독립변수의 관계식은?

→ $Y = \beta_0 + \beta_1 X + \varepsilon$

36.회귀분석에서 오차항 ε 의 의미는?

→ 설명되지 않는 변동(잡음)

37.최소제곱법(OLS)은 무엇을 최소화?

→ 잔차제곱합

38.회귀모형에서 결정계수 R^2 는 무엇을 의미?

→ 설명력 (0~1 사이)

39.다중회귀에서 다중공선성 문제란?

→ 독립변수들 간 높은 상관

40.회귀분석 가정 4가지 중 하나를 쓰시오.

→ (예: 오차의 독립성, 등분산성, 정규성, 선형성)

⑤ 기타 통계 개념

41. 표본과 모집단의 차이는?

→ 모집단: 전체 / 표본: 일부

42. 표본추출 방법 중 단순무작위추출이란?

→ 모든 개체가 동일 확률로 선택

43. 층화추출의 장점은?

→ 집단별 대표성 보장, 추정의 정확성 증가

44. 통계학에서 모수(parameter)란?

→ 모집단을 설명하는 값(평균, 분산 등)

45. 통계량(statistic)이란?

→ 표본으로 계산한 값

46. 편향(bias)이란?

→ 추정치가 체계적으로 실제 모수와 차이남

47. 표본수가 커질수록 추정량이 안정적이라는 성질은?

→ 일치성(consistency)

48. 분산이 크다는 의미는?

→ 데이터의 흩어짐이 크다

49. 중앙값이 평균보다 크다는 것은 어떤 분포 모양?

→ 왼쪽으로 꼬리(좌측 비대칭)

50. 히스토그램은 어떤 목적으로 사용하는가?

→ 데이터의 분포 형태 시각화