

21.1 Conditioning, Bayesian inference and sampling error

Name: _____ Class: _____ Date: _____ **Total: 10 marks**

KEY WORDS posterior probability • standard error

Objective

Separate conditional, joint and independent-event probabilities; Calculate posterior probabilities with a complete base-rate table; Use variance and sample size to determine a sample mean standard error.

Original public practice. Marking guidance is authored locally. Use the topic reference for prerequisites and full theorem statements.

1 Key knowledge and worked examples

posterior probability : A probability conditional on the observed evidence after accounting for prior proportions.

standard error : The standard deviation of a statistic across repeated samples.

For $P(B)>0$, $P(A|B)=P(A\cap B)/P(B)$: restrict the population to B before computing the fraction in A. Independence means $P(A\cap B)=P(A)P(B)$, equivalently $P(A|B)=P(A)$ when the denominator is nonzero. Mutually exclusive events with positive probabilities cannot be independent, because their joint probability is zero. Sampling without replacement usually changes later probabilities; a fixed denominator for successive draws would describe a different experiment.

Let $P(A)=0.4$, $P(B)=0.5$ and $P(A\cap B)=0.3$. Then $P(A|B)=0.6$, so A and B are not independent, since $0.3\neq 0.4 \cdot 0.5$. If a variable has standard deviation 12, an independent sample of size 36 has mean standard error 2. Increasing the sample size to 144 gives 1; the individual-observation standard deviation stays 12.

Condition on positives: 80 / 125, not 80 / 100

Expected group	Positive	Negative	Total
Defective	80	20	100
Not defective	45	855	900
Total	125	875	1000

Given a positive test, divide by all 125 positives: 80/125, not 80/100

2 Practice

2.1 $P(A)=0.4$, $P(B)=0.5$ and $P(A\cap B)=0.3$. What is $P(A|B)$? [1]

- A 0.2
- B 0.3
- C 0.4
- D 0.5
- E 0.6

2.2 An iid sample of size 25 has mean standard error 6. What is the mean standard error for size 100 from the same population? [1]

- A 1.5
- B 3
- C 6
- D 12
- E 24

2.3 In a stated model, 10% of devices have a defect. A test is positive for 80% of defective and 5% of nondefective devices. Find the probability of a defect given a positive result using 1000 expected devices. ^[4]

3 Exam-style questions

3.1 Independent observations have variance 25. Find the standard error of their mean for $n=100$. If instead every distinct pair has covariance 1, find the mean's variance. Does quadrupling n necessarily halve the standard error in that second model? ^[4]

Solutions

2.1 E. Restrict to B: divide the joint probability 0.3 by $P(B)=0.5$, obtaining 0.6. The conditioned denominator is 0.5. The smaller numerical choices use a marginal, joint probability or different operation, and none equals 0.3 divided by 0.5.

2.2 B. Population standard deviation is $6\sqrt{25}=30$. Dividing by $\sqrt{100}=10$ gives standard error 3. Quadrupling sample size divides standard error by two. One-point-five divides by four; six leaves it fixed; twelve and twenty-four reverse the direction.

2.3 Expected counts are 100 defective and 900 nondefective. Positive counts are $100(0.8) = 80$ and $900(0.05) = 45$, respectively. The conditioned group contains 125 positives, hence

$$P(D \mid +) = \frac{80}{80 + 45} = \frac{16}{25} = 0.64.$$

Negative counts 20 and 855 check both row totals. The sensitivity 0.8 is $P(+ \mid D)$, not the requested reversed condition. These are illustrative model rates, not measured performance claims.

3.1 Independence gives $\text{Var}(\bar{X}) = 25/n$ and at 100 standard error $5/10 = 0.5$. With covariance one,

$$\text{Var}(\bar{X}) = \frac{25n + n(n-1)}{n^2} = 1 + \frac{24}{n}.$$

At 100 it is 1.24. The standard error tends to one as n grows, rather than to zero; quadrupling n need not halve it. Covariance terms may be dropped only under justified uncorrelatedness, with independence a sufficient condition.

Local four-mark guidance for each written task: choose and justify the method; carry through the reasoning; give every requested result; check the relevant restriction, hypothesis or boundary case. These are classroom marks, not GRE scoring.