T.8 · Conditioning, Bayesian inference and sampling error

Original classroom teaching; not an official form or score predictor

Conditioning, Bayesian inference and sampling error

- Compute conditional probabilities and use the law of total probability.
- Find a **posterior probability** 后验概率 with Bayes' rule.
- Separate independence from mutual exclusion.
- Use expectation and variance of sums, and the **standard error** 标准误 of a mean.

Conditional probability

- The second formula, the law of total probability, needs the A_i to split the whole space.
- $P(A \mid B)$ and $P(B \mid A)$ are usually different.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) = \sum_i P(B \mid A_i) P(A_i)$$

Bayes' rule

- The prior $P(A)$ is the starting probability; the posterior $P(A \mid B)$ is updated by the evidence.
- Expected counts make it concrete: imagine 1,000 cases and fill in a table.
- A rare cause stays fairly unlikely after one positive test unless the test rarely gives false positives.

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Figure: Bayes' rule

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

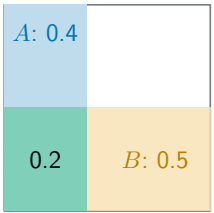
Expected group	Positive	Negative	Total
Defective	80	20	100
Not defective	45	855	900
Total	125	875	1000

- Given a positive test, divide by all 125 positives: 80/125, not 80/100
- Match each part of the diagram to the text.

Independent and exclusive

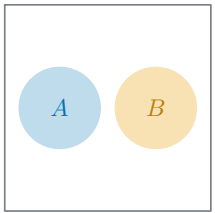
- Independent: $P(A \cap B) = P(A)P(B)$. Knowing A tells you nothing about B.
- Mutually exclusive: $P(A \cap B) = 0$.
- Two events with positive probabilities cannot be both: if they are exclusive, knowing A tells you B did not happen.

Figure: Independent and exclusive



independent

$P(A \cap B) = 0.4 \times 0.5$



mutually exclusive

$P(A \cap B) = 0 \neq P(A)P(B)$

- Independent is not the same as mutually exclusive.
- Match each part of the diagram to the text.

Sums and sampling error

- $E[aX + bY] = aE[X] + bE[Y]$ always. $\mathrm{Var}(X + Y) = \mathrm{Var}\, X + \mathrm{Var}\, Y$ when X and Y are independent.
- The mean of n independent observations with standard deviation σ has standard error $\sigma/\sqrt{n}$.
- Central limit theorem: for large n the sample mean is approximately normal. A 95% interval is about the estimate ± 1.96 standard errors.
- To halve the standard error, quadruple the sample.

Worked example: a defect test

- In 1,000 devices: 100 defective, 80 positive; 900 good, 45 positive.
- $P(\text{defect} \mid +) = \dfrac{80}{80+45} = \dfrac{16}{25} = 0.64.$
- The 0.8 is $P(+ \mid \text{defect})$, the reversed condition. Check: the negatives are $20 + 855 = 875$, and $125 + 875 = 1000$.

Worked example: sampling error

- Standard error: $10/\sqrt{25} = 2$.
- An approximate 95% interval is $72 \pm 1.96 \times 2 = 72 \pm 3.9$.
- With 100 students the standard error would be 1: four times the sample for half the error.

Common errors

- Confusing $P(A \mid B)$ with $P(B \mid A)$.
- Ignoring the prior (base rate) in Bayes' rule.
- Treating exclusive events as independent.

Common errors (2)

- Adding standard deviations instead of variances.
- Thinking the standard error falls in proportion to n.

Key terms

- **posterior probability**: The probability of a cause after taking account of the observed evidence.
- **standard error**: The standard deviation of a statistic, such as a sample mean, across repeated samples.

Check yourself

- A and B are independent with $P(A) = 0.3$, $P(B) = 0.5$. What is $P(A \cup B)$?
- A family has two children, at least one a girl. What is the probability both are girls?
- Observations have SD 8. What is the standard error of the mean of 16 of them?