

# Probability & inference

10 essentials · conditions, examples & common traps

---

## 01. Conditional probability

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

Use when B is known; require  $P(B) > 0$ . The denominator is the group you are already inside.

## 02. Bayes & total probability

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

For a partition  $A_i$ ,  $P(B) = \sum P(B | A_i)P(A_i)$ . Count every route into B; a rare cause can stay unlikely.

## 03. Independence vs. disjoint events

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

Independence means knowing one does not change the other. Disjoint events cannot both occur; positive-probability disjoint events are not independent.

## 04. Expected value & variance

$$E[X] = \sum_x xp(x), \quad \text{Var}(X) = E[X^2] - E[X]^2$$

For discrete X, weight each value by its probability.  $E[aX+b] = aE[X]+b$ ;  $\text{Var}(aX+b) = a^2\text{Var}(X)$ .

## 05. Adding random variables

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

Expectations always add when finite. Variances add without the covariance term only when covariance is zero; independence is sufficient.

## WORKED EXAMPLE 1

**Q.** 15 of 30 Statistics students also take Linear Algebra. Find  $P(\text{Linear Algebra} | \text{Statistics})$ .

**Answer.** 0.50. Restrict to the 30 Statistics students:  $15/30 = 0.50$ . The full campus is not the denominator.

**Variation.** If 10 of 50 take both, given Statistics:  $10/50 = 0.20$ . Same rule, different group size.

## WORKED EXAMPLE 3

**Q.** A test has 90% sensitivity, 5% false positives, and 1% prevalence. After a positive result, what is  $P(\text{disease})$ ?

**Answer.** 15.4%. Use  $0.009 / (0.009 + 0.0495)$ . Most people are healthy, so false positives matter.

**Variation.** At 2% prevalence, the result is  $0.018 / 0.067 = 26.9\%$ .

**CHOOSE THE METHOD** Choose the model first: a count of independent successes suggests binomial; a sample mean needs a standard error. Report units and the population you are describing.

**WATCH OUT** Do not swap  $P(A | B)$  with  $P(B | A)$ . SD describes observations; SE describes an estimator. Check independence before adding variances.

## FINAL CHECK

For  $P(A | B)$ , divide by  $P(B)$ . First name the group that is already known.

A larger n reduces the standard error, not the spread of individual observations.

State the model assumptions and units. A small p-value does not measure effect size.

## 06. Binomial model

$$X \sim \text{Bin}(n, p) : \quad \mu = np, \quad \sigma^2 = np(1-p)$$

Count successes in n independent trials with the same success probability.  $P(X=k) = C(n,k)p^k(1-p)^{n-k}$ .

## 07. Normal scores & sample means

$$z = \frac{x - \mu}{\sigma}, \quad \text{SE}(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

The standard error formula needs independent observations with common finite variance. A large-sample normal approximation also needs a suitable CLT.

## 08. Confidence interval for a mean

$$\bar{x} \pm t_{n-1}^* \frac{s}{\sqrt{n}}$$

Use the t method for independent normal data; large samples often permit approximation. A 95% method covers the fixed mean in 95% of repeated samples.

## 09. Hypothesis tests

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad p \leq \alpha \Rightarrow \text{reject } H_0$$

Choose the null, alternative and  $\alpha$  first. A p-value measures how extreme the data are under  $H_0$ ; it is not  $P(H_0 \text{ is true})$ .

## 10. Errors & power

$$\text{Type I: reject true } H_0, \quad \text{power} = 1 - \beta$$

Type II means failing to reject a false null. Failing to reject does not prove no effect. Statistical significance does not measure effect size.

## WORKED EXAMPLE 2

**Q.** A fair coin is tossed 10 independent times. Find the mean and variance of the number of heads.

**Answer.** Mean 5; variance 2.5. Here  $n=10$ ,  $p=0.5$ , so  $np=5$  and  $np(1-p)=2.5$ .

**Variation.** For 20 trials with  $p=0.2$ : mean 4, variance 3.2. Use  $20 \times 0.2 \times 0.8$  for the variance.

## WORKED EXAMPLE 4

**Q.** For independent normal observations,  $n=25$ , mean=20,  $s=5$ . Use  $t^*=2.064$  for a 95% mean interval.

**Answer.** [17.94, 22.06].  $\text{SE}=5/\sqrt{25}=1$ , so take  $20 \pm 2.064$ .

**Variation.** With  $n=100$  and the same s, SE falls to 0.5. Re-compute the critical value for a new interval.