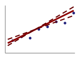
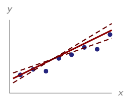
Inference for Slopes

AP Statistics ▪ Unit 9

AP Statistics



What we will cover



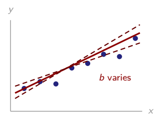
1 Slope

The slope is a statistic

A regression line comes from *one* sample, so the sample **slope** 斜率 b has a **sampling distribution** 抽样分布.

- true population slope β is fixed but unknown
- b estimates β, bouncing sample to sample

> Inference measures the **uncertainty** 不确定性 in b to speak about the true relationship.



2 Interval

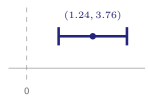
Confidence interval for a slope

Build a t-interval on the true slope β. Check the conditions – **LINER**: **L**inear, **I**ndependent, **N**ormal, **E**qual spread, **R**andom.

$$b \pm t^* \cdot SE_b$$

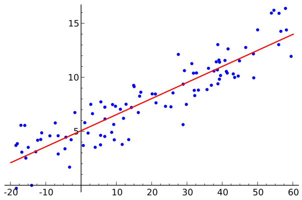
- read b and SE_b from computer output
- $df = n - 2$ (a line uses a and b)

> If the interval excludes 0, there is a real linear relationship.

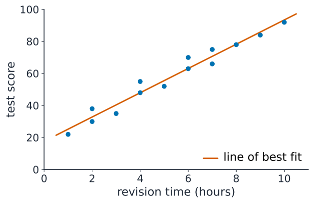


$b = 2.5,\ SE_b = 0.6,\ df = 18.$

Least-squares regression: the line that minimises



Least-squares regression: the line that minimises the sum of squared residuals



Slope inference is based on the least-squares regression line through the points

> Least-squares regression: the line that minimises the sum of squared residuals

3 Test

Test for a slope

The test asks directly whether the slope is zero:

- **null** 原假设 $H_0 : \beta = 0$ (no linear relationship)
- **alternative** 备择假设 $H_a : \beta >, <,$ or $\neq 0$

$$t = \frac{b}{SE_b}, \qquad df = n - 2$$

> $\beta = 0$ means a flat line – x has no straight-line effect on y.

Study time

$b = 2.5,\ SE_b = 0.6,\ n = 20$:

$$t = \frac{2.5}{0.6} \approx 4.17$$

$df = 18$, p ≈ 0.0006: reject – a real linear link.

4 Select

Selecting the right procedure

The whole course in one decision – read the type of data first:

Proportions categorical: one/two-sample z	**Means** quantitative: one/two-sample or paired t
Count tables χ^2: fit, homogeneity, independence	**Slope** two quantitative vars: t for β

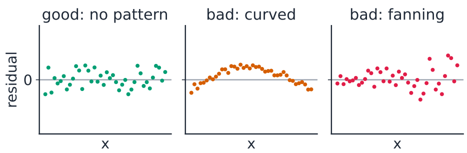
> **Select** 选择 the procedure, **implement** 实施 with conditions, **communicate** 表达 in context.

Choosing the right inference procedure

1. what is estimated	a proportion, a mean, a difference, a distribution of counts, or a slope
2. how many samples	one, two independent, or paired
3. interval or test	an interval estimates a value; a test judges a claim
4. which model	z for proportions, t for means, χ^2 for counts
5. the conditions	random, independent (10% rule), and normal or large enough
then	name the procedure explicitly before any calculation

> Naming the procedure and checking its conditions is worth marks before any arithmetic. On the free-response paper it is often the larger half.

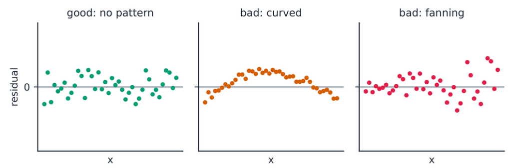
Confidence Interval for a Slope, continued



residual $= y - \hat{y}$ · random scatter about 0 supports a linear model (LINER)

> A random, patternless residual plot supports the conditions; a curve or a fan does not

Setting Up a Test for a Slope



residual $= y - \hat{y}$ · random scatter about 0 supports a linear model (LINER)

> Check residual plots before trusting a slope CI or test

Do Those Points Align?

A sample **scatterplot 散点图** gives a **sample slope 样本斜率** *b* for the least-squares **regression 回归** line – but a different sample would give a slightly different slope.

> So *b* is a statistic with **sampling variability 抽样变异性**, estimating the **true (population) slope 总体斜率** β.

Exam tips

- Inference for a **slope** tests whether the true slope is 0 (no linear relationship).
- If a slope's confidence interval **includes 0**, you cannot conclude a real **linear** relationship – the variables may still be related in a curved way.
- Read the slope, standard error, t-statistic, and p-value straight from computer output – but the printed p-value is **two-tailed**, so halve it for a one-tailed H_a.
- Check the regression conditions (linearity, independence, roughly normal residuals, equal spread) via the residual plot.
- Interpret the interval and test in context, tied to the true slope.

5 Recap

Unit 9 – key takeaways

1 Slope varies
b estimates the true slope β; it has a sampling distribution

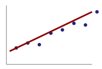
2 Interval
$b \pm t^* SE_b$, $df = n - 2$; LINER conditions

3 Test
$H_0 : \beta = 0$; $t = b/SE_b$; excludes 0 $\Rightarrow$ real link

4 Choose
data type picks z, t, χ^2, or slope-t

Check yourself

1. What are the degrees of freedom for a slope t-procedure?
2. What does the LINER "E" stand for?
3. What null value of β means no linear relationship?
4. Two quantitative variables and a relationship: which test?



Answers **1.** $n - 2$ **2.** Equal spread **3.** $\beta = 0$ **4.** a t-test for the slope