

Exercise walk-through

4.2

Inference using normal and t-distributions

A-Level Further Mathematics

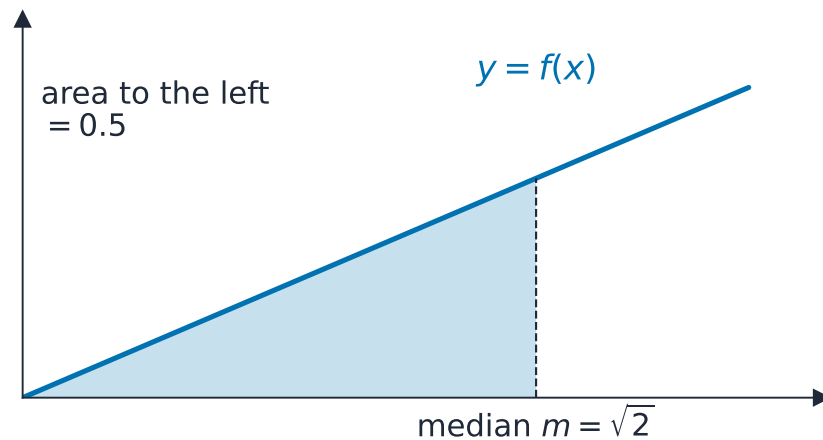
Continuous random variables

$$F(x) = P(X \leq x) = \int_{-\infty}^x f, \text{ and } f = F'.$$

$$f(x) = \frac{1}{2}x, 0 \leq x \leq 2$$

$$F = \frac{1}{4}x^2; \text{ median } m = \sqrt{2} = 1.41$$

Continuous random variables, in detail



$$\text{median } m \text{ solves } F(m) = \int_{-\infty}^m f(x) dx = \frac{1}{2}$$

median: area 0.5 to its left

$$F(x) = P(X \leq x) = \int_{-\infty}^x f, \text{ and } f = F'.$$

The cumulative distribution function

cdf

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt - \text{the running total of the density}$$

going back

$$f(x) = \frac{dF}{dx}: \text{differentiate the cdf to recover the density}$$

its properties

$$\text{non-decreasing, } F(-\infty) = 0, F(\infty) = 1$$

the median

$$\text{solve } F(m) = 0.5; \text{ percentiles}$$

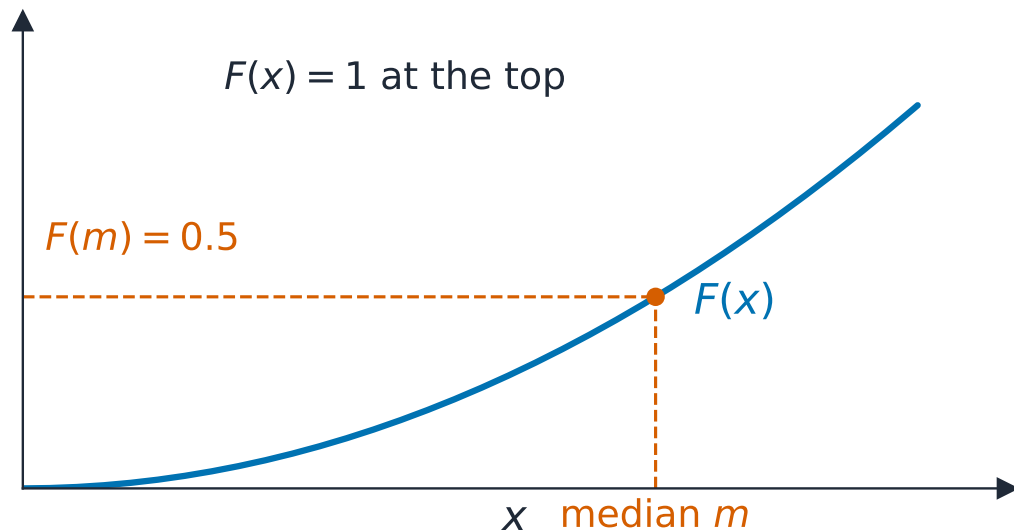
$$\text{solve } F(x) = p$$

why it matters

$$\text{any probability is a difference of two values: } P(a < X \leq b) = F(b) - F(a)$$

For a probability, reach for F ; for a mode or an expectation, reach for f . Knowing which one the question needs saves the whole calculation.

The cumulative graph $F(x)$ rises from 0



$$F(x) = P(X \leq x) = \int_{-\infty}^x f \cdot 0 \leq F \leq 1 \cdot \text{median: } F(m) = \frac{1}{2}$$

The cumulative graph $F(x)$ rises from 0 to 1; the height 0.5 is reached at the median.

The distribution of a related variable

the setup

X has a known distribution and $Y = g(X)$; find Y 's

the method

work through the **cumulative** function, never the density directly

for $Y = X^2$

$$F_Y(y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

then

differentiate to get $f_Y(y)$

the caution

mind the **range** of Y , and whether g is one-to-one over X 's range

Never substitute into the density and hope. The cdf route is the only one that handles a many-to-one g correctly.

Exercises 2.1·2.2

2.1 State the number of degrees of freedom for a one-sample t-test with $n = 12$. [1]

2.2 State when a t-distribution is used instead of the normal. [1]

Exercises 2.1·2.2

2.1 State the number of degrees of freedom for a one-sample t-test with $n = 12$. [1]

2.1

- $n - 1 = 11$

2.2 State when a t-distribution is used instead of the normal. [1]

2.2

- when the sample is small and the population variance is unknown

Exercise 3.1

3.1 A 95% confidence interval based on a small sample uses a t value that is, compared with 1.96: [1]

☐

A smaller

B larger

C equal

D zero

Exercise 3.1

3.1 A 95% confidence interval based on a small sample uses a t value that is, compared with 1.96: [1]

☐

A smaller

B larger

C equal

D zero

3.1 B

- The small-sample t value is larger than 1.96

Exercise 3.2 (a)

3.2 A sample of 8 measurements has $\bar{x} = 20.5$ and $s = 1.2$.

(a) Find a 95% confidence interval for the population mean (use $t = 2.365$, 7 d.f.). [3]

Exercise 3.2 (a)

3.2 A sample of 8 measurements has $\bar{x} = 20.5$ and $s = 1.2$.

(a) Find a 95% confidence interval for the population mean (use $t = 2.365$, 7 d.f.). [3]

(a)

- $20.5 \pm 2.365 \frac{1.2}{\sqrt{8}} = 20.5 \pm 2.365(0.4243) = 20.5 \pm 1.00; (19.5, 21.5)$

Exercise 3.2 (b)

3.2 A sample of 8 measurements has $\bar{x} = 20.5$ and $s = 1.2$.

(b) State one assumption needed for this interval to be valid.

[1]

Exercise 3.2 (b)

3.2 A sample of 8 measurements has $\bar{x} = 20.5$ and $s = 1.2$.

(b) State one assumption needed for this interval to be valid.

[1]

(b)

- the population is (approximately) normally distributed

Exercise 3.3

3.3 A machine should produce rods of mean length 25 cm. A sample of 6 rods has $\bar{x} = 25.4$ cm and $s = 0.5$ cm. Test at the 5% level (two-tailed, $t_{\text{crit}} = 2.571$) whether the mean differs from 25. [5]

Exercise 3.3

3.3 A machine should produce rods of mean length 25 cm. A sample of 6 rods has $\bar{x} = 25.4$ cm and $s = 0.5$ cm. Test at the 5% level (two-tailed, $t_{\text{crit}} = 2.571$) whether the mean differs from 25. [5]

3.3

- $H_0: \mu = 25$, $H_1: \mu \neq 25$; $t = \frac{25.4 - 25}{0.5/\sqrt{6}} = \frac{0.4}{0.2041} = 1.96$; compare with 2.571; since $1.96 < 2.571$, do not reject H_0 : no evidence the mean differs from 25