

# QF632-2025-W3

Number of participants: 40

1. Suppose  $X$  has PDF  $f(x) = 3x^2$ ,  $0 \leq x \leq 1$  and zero elsewhere. What is  $P(X = 0.5)$ ?
- 8 correct answers  
out of 31  
respondents

$f(0.5) = 3 \cdot (0.5)^2$



15 votes

$\int_{0.5}^{0.5} 3x^2 dx$



6 votes



0



8 votes

$P(0 \leq X \leq 0.5)$



2 votes

Let  $X$  be a continuous random variable with CDF  $F$  and PDF

17 correct answers

out of 31 respondents



2.  $f$ . Which of the following statements must hold (wherever the derivatives exist)?

$$F(x) = \int_x^{\infty} f(t) dt$$



8 votes



$$f(x) = \frac{d}{dx} F(x)$$



17 votes

$$F(x) = f(x)$$



1 vote

$$\int_{-\infty}^x f(t) dt = 1 - F(x)$$



5 votes



Let  $Y = aX + b$ , where  $a, b$   
3. are constants and  $X \sim \mathcal{N}(\mu, \sigma^2)$ . What is  $\text{Var}(Y)$ ?

24 correct  
answers  
out of 31  
respondents



$$a^2 \sigma^2$$



24 votes

$$a \sigma^2$$



1 vote

$$\sigma^2$$



6 votes

$$b^2 \sigma^2$$



0 votes



#### Which of the following 4. distance measures does OLS minimize?

29 correct  
answers  
out of 32  
respondents



Sum of squared  
vertical distances  
 $(y_i - \hat{y}_i)^2$



91%

29 votes

Sum of squared  
perpendicular  
distances to the  
hyperplane



6%

2 votes

Maximum  
absolute  
deviation from  
the hyperplane



0%

0 votes

Sum of absolute  
vertical distances  
 $|y_i - \hat{y}_i|$



3%

1 vote



**5. In the fitted model  $\hat{y} = \beta_0 + \sum_{j=1}^p \beta_j x_j$ , the coefficient  $\beta_k$  is interpreted as:**

**26 correct answers**  
out of 27 respondents

The change in  $\hat{y}$  when  $\textit{all}$  predictors increase by one unit.

0%

0 votes

The fraction of total variance in  $y$  explained by  $x_k$  alone.

4%

1 vote

The change in  $\hat{y}$  when  $x_k$  increases by one unit, holding all other  $x_j$  fixed.

96%

26 votes

The correlation between  $x_k$  and the residuals.

0%

0 votes

Consider the following causal diagram:  $Z \longrightarrow X \longrightarrow$

$Y, \quad Z \longrightarrow Y$ . Which of the

following is **true** about the marginal correlation  $\text{Corr}(X, Y)$  vs.

the partial correlation

$\text{Corr}(X, Y \mid Z)$ ?

24 correct answers

out of 32 respondents



6.

They must always agree in sign.



4 votes



They can disagree in sign if  $Z$  is a confounder.



24 votes

Marginal correlation is zero if and only if the partial correlation is zero.



2 votes

Partial correlation can only be larger in magnitude than the marginal correlation.



2 votes

Consider the model  $mpg_i = \beta_0 + \beta_1 qsec_i + \beta_2 am_i + \beta_3 (qsec_i \times am_i) + \varepsilon_i$ , where



7.  $am_i = 0$  for automatic and  $am_i = 1$  for manual. Which of the following is the slope of  $qsec$  for manual cars?

26 correct answers

out of 29 respondents

$\beta_1$

0%

0 votes

$\beta_2$

0%

0 votes



$\beta_1 + \beta_3$

90%

26 votes

$\beta_3$

10%

3 votes



8. At what value of  $q_{sec}$  do the two fitted lines  $(\beta_0 + \beta_1 q$  and  $\beta_0 + \beta_2 + (\beta_1 + \beta_3)q)$  intersect?

19 correct answers  
out of 25 respondents



$$q = -\frac{\beta_2}{\beta_3}$$



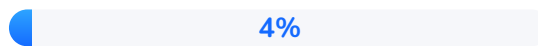
19 votes

$$q = \frac{\beta_2}{\beta_1}$$



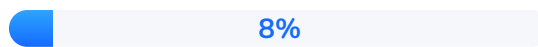
3 votes

$$q = -\frac{\beta_1}{\beta_2}$$



1 vote

$$q = \frac{\beta_3}{\beta_2}$$



2 votes

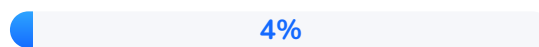




9. You fit the models (A)  $y = \beta_0 + \beta_1 x + \varepsilon$  and (B)  $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$ . Which formal test assesses whether the quadratic term adds significant explanatory power?

17 correct answers  
out of 28 respondents

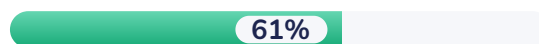
Test  $H_0: \beta_1 = 0$  under model (B)



1 vote



Test  $H_0: \beta_2 = 0$  in model (B)



17 votes

Compare  $R^2$  of (A) and (B) without an  $F$ -test



7 votes

Test  $H_0: \beta_0 = \beta_1 = \beta_2 = 0$



3 votes

## The least-squares objective

$$S(\beta) = (y - X\beta)^\top (y -$$



10.  $X\beta)$  has Hessian  $2X^\top X$ .

Which property of  $X^\top X$  ensures  $S(\beta)$  is convex?

19 correct answers

out of 24 respondents

$X^\top X$  is diagonal.



17%

4 votes



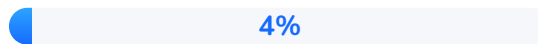
$X^\top X$  is positive semidefinite.



79%

19 votes

$X^\top X$  has full row rank.



4%

1 vote

$\det(X^\top X) = 0$



0%

0 votes



11.

Consider the ridge loss  $L_{\text{ridge}} = \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2$ . As  $\lambda \rightarrow \infty$ , the ridge estimates  $\hat{\beta}_j^{\text{ridge}}$  for  $j = 1, \dots, p$  tend to:

13 correct answers  
out of 25 respondents

the ordinary-least-squares estimates.



5 votes

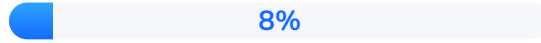


0 (while  $\beta_0$  remains the OLS intercept).



13 votes

infinity in magnitude.



2 votes

random values with mean zero.



5 votes



## 12. Why is it essential to standardize (center and scale) predictors before fitting a lasso model?

**25 correct answers**  
out of 28 respondents

The lasso penalty  $\sum_j |\beta_j|$  is invariant to predictor scaling.



3 votes



Without standardization, predictors with larger scale incur disproportionately larger penalties.



25 votes

Standardization guarantees a unique solution.



0 votes

Only binary predictors must be standardized for lasso.



0 votes



**13. In the high-dimensional regime where  $p > n$ , a Lasso solution can have at most how many nonzero coefficients?**

**17 correct answers**

out of 27 respondents



$n$



63%

17 votes

$p$



19%

5 votes

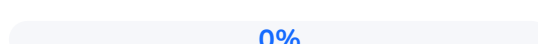
$\sqrt{n}$



19%

5 votes

$\log(p)$



0%

0 votes