PCA and Exploratory Factor Analysis – Quiz

Statistical Data Analysis

Question 1

PCA transforms original variables into new variables called principal components. What geometric property do these components satisfy?

A. They are orthogonal to each other and ordered by decreasing variance

B. They are parallel to the original axes

C. They are aligned with the variables that have the largest mean

D. They always form a 45-degree angle with the first original variable

Question 1

PCA transforms original variables into new variables called principal components. What geometric property do these components satisfy?

A. They are orthogonal to each other and ordered by decreasing variance

B. They are parallel to the original axes

C. They are aligned with the variables that have the largest mean

D. They always form a 45-degree angle with the first original variable

Answer: A

Principal components are mutually orthogonal (perpendicular) directions in the data space, ordered so that PC1 captures the most variance, PC2 the second most, and so on. This orthogonality ensures the components are uncorrelated, which is a defining property of PCA.

Question 2

A researcher has 12 highly correlated sensor measurements and wants to visualize the data in two dimensions for exploratory analysis. No theoretical model exists for the sensors. Which method is most appropriate?

A. Confirmatory Factor Analysis (CFA)

B. Exploratory Factor Analysis with Promax rotation

C. Multiple linear regression with 12 predictors

D. PCA retaining the first two components

Question 2

A researcher has 12 highly correlated sensor measurements and wants to visualize the data in two dimensions for exploratory analysis. No theoretical model exists for the sensors. Which method is most appropriate?

A. Confirmatory Factor Analysis (CFA)

B. Exploratory Factor Analysis with Promax rotation

C. Multiple linear regression with 12 predictors

D. PCA retaining the first two components

Answer: D

PCA is the standard choice for dimensionality reduction and visualization when no theoretical model about latent constructs exists. Retaining the first two PCs provides the best two-dimensional summary of the 12-variable data by maximizing the variance captured in two dimensions.

Question 3

The covariance matrix Σ of three standardized variables is a 3×3 matrix. What values appear on its diagonal when the analysis is based on the correlation matrix?

A. The means of each variable

B. All ones, because each standardized variable has variance 1

C. The eigenvalues of each component

D. The communalities of each variable

Question 3

The covariance matrix Σ of three standardized variables is a 3×3 matrix. What values appear on its diagonal when the analysis is based on the correlation matrix?

A. The means of each variable
B. All ones, because each standardized variable has variance 1
C. The eigenvalues of each component
D. The communalities of each variable

Answer: B

When variables are standardized (mean 0, standard deviation 1), the covariance matrix equals the correlation matrix. The diagonal entries are the variances, which are all 1 for standardized variables. This is why the Kaiser criterion uses the threshold of 1: an eigenvalue above 1 captures more variance than any single standardized variable contributes.

Question 4

Given a 2×2 covariance matrix $\Sigma = \begin{pmatrix} 4 & 2 2 & 3 \end{pmatrix}$, what is the sum of its eigenvalues?

A. 5

B. 6

C. 7

D. 8

Question 4

Given a 2×2 covariance matrix $\Sigma = \begin{pmatrix}4 & 2\, 2 & 3\end{pmatrix}$, what is the sum of its eigenvalues?

A. 5

B. 6

C. 7

D. 8

Answer: C

The sum of eigenvalues of any square matrix equals its trace (sum of diagonal elements). Here, $\text{trace}(\Sigma) = 4 + 3 = 7$. This property is fundamental in PCA because it means the total variance in the data is preserved and merely redistributed across the principal components.

Question 5

For the eigenvalue equation $\Sigma v = \lambda v$, which mathematical condition must be satisfied to find the eigenvalues λ?

A. The trace of Σ must equal zero

B. $\det(\Sigma - \lambda I) = 0$

C. The inverse of $\Sigma - \lambda I$ must exist

D. The matrix Σ must be invertible

Question 5

For the eigenvalue equation $\Sigma v = \lambda v$, which mathematical condition must be satisfied to find the eigenvalues λ?

A. The trace of Σ must equal zero

B. $\det(\Sigma - \lambda I) = 0$

C. The inverse of $\Sigma - \lambda I$ must exist

D. The matrix Σ must be invertible

Answer: B

Eigenvalues are found by solving the characteristic equation $\det(\Sigma - \lambda I) = 0$. This determinant condition ensures that $(\Sigma - \lambda I)v = 0$ has a non-trivial solution $v \neq 0$. For a $p \times p$ matrix, this yields a polynomial of degree p whose roots are the eigenvalues.

Question 6

A PCA on 7 variables yields eigenvalues 3.2, 1.8, 0.9, 0.5, 0.3, 0.2, 0.1. What percentage of total variance does PC1 explain?

A. Approximately 46%
B. Approximately 32%
C. Approximately 55%
D. Approximately 71%

Question 6

A PCA on 7 variables yields eigenvalues 3.2, 1.8, 0.9, 0.5, 0.3, 0.2, 0.1. What percentage of total variance does PC1 explain?

A. Approximately 46%

B. Approximately 32%

C. Approximately 55%

D. Approximately 71%

Answer: A

The total variance equals the sum of all eigenvalues: $3.2 + 1.8 + 0.9 + 0.5 + 0.3 + 0.2 + 0.1 = 7.0$. The proportion explained by PC1 is $3.2 / 7.0 = 0.457$, or approximately 46%. This single component captures nearly half of the information in all seven original variables.

Question 7

In a scree plot, a researcher observes that eigenvalues drop sharply from component 1 to 3, then level off from component 4 onward. Using the elbow method, how many components should be retained?

A. 1 component

B. 2 components

C. 3 components

D. 4 components

Question 7

In a scree plot, a researcher observes that eigenvalues drop sharply from component 1 to 3, then level off from component 4 onward. Using the elbow method, how many components should be retained?

A. 1 component

B. 2 components

C. 3 components

D. 4 components

Answer: C

The elbow method retains the components before the point where the curve flattens out. If the curve levels off starting at component 4, the elbow is at component 3, so components 1 through 3 should be retained. These three components capture the meaningful variance before diminishing returns set in.

Question 8

A PCA on 6 standardized variables yields eigenvalues 2.9, 1.4, 1.05, 0.85, 0.45, 0.35. Applying the Kaiser criterion, how many components are retained, and what is a known limitation of this rule?

A. 2 components; it tends to retain too few components

B. 4 components; it always gives the optimal number

C. 3 components; it tends to retain too few components

D. 3 components; it tends to retain too many components with large variable sets

Question 8

A PCA on 6 standardized variables yields eigenvalues 2.9, 1.4, 1.05, 0.85, 0.45, 0.35. Applying the Kaiser criterion, how many components are retained, and what is a known limitation of this rule?

A. 2 components; it tends to retain too few components

B. 4 components; it always gives the optimal number

C. 3 components; it tends to retain too few components

D. 3 components; it tends to retain too many components with large variable sets

Answer: D

The Kaiser criterion retains components with eigenvalues greater than 1. Here, three eigenvalues exceed 1 (2.9, 1.4, 1.05), so three components are kept. A known limitation, as taught in the lesson, is that this rule tends to over-extract components, especially when the number of variables is large.

Question 9

Parallel analysis determines the number of factors by comparing observed eigenvalues to eigenvalues obtained from:

A. A theoretical chi-square distribution

B. Random data with the same number of variables and observations

C. The inverse of the covariance matrix

D. A bootstrap sample of the original dataset

Question 9

Parallel analysis determines the number of factors by comparing observed eigenvalues to eigenvalues obtained from:

A. A theoretical chi-square distribution

B. Random data with the same number of variables and observations

C. The inverse of the covariance matrix

D. A bootstrap sample of the original dataset

Answer: B

Parallel analysis generates many random datasets with the same dimensions as the real data, computes eigenvalues for each, and averages them. Only factors whose observed eigenvalues exceed the corresponding random-data eigenvalues are retained. The lesson describes this as the gold standard for determining the number of factors.

Question 10

In a PCA biplot, two variable arrows point in nearly the same direction. What does this indicate about those two variables?

A. They are negatively correlated with each other

B. They are uncorrelated with each other

C. They have low communalities

D. They are strongly positively correlated with each other

Question 10

In a PCA biplot, two variable arrows point in nearly the same direction. What does this indicate about those two variables?

A. They are negatively correlated with each other

B. They are uncorrelated with each other

C. They have low communalities

D. They are strongly positively correlated with each other

Answer: D

In a biplot, the angle between variable arrows reflects the correlation between those variables. An angle near 0 degrees indicates strong positive correlation, 90 degrees indicates near-zero correlation, and 180 degrees indicates strong negative correlation. Two arrows pointing in the same direction thus reveal a strong positive relationship.

Question 11

A biplot shows five variable arrows. Three arrows (Income, Education, OccupationLevel) point strongly to the right along PC1, while two arrows (Age, Tenure) point strongly upward along PC2. A data point appears in the upper-right quadrant. What can be inferred?

A. The observation scores low on both PC1 and PC2

B. The observation has high values on Income, Education, OccupationLevel but low Age and Tenure

C. The observation is an outlier that must be removed

D. The observation has high values on all five variables

Question 11

A biplot shows five variable arrows. Three arrows (Income, Education, OccupationLevel) point strongly to the right along PC1, while two arrows (Age, Tenure) point strongly upward along PC2. A data point appears in the upper-right quadrant. What can be inferred?

A. The observation scores low on both PC1 and PC2

B. The observation has high values on Income, Education, OccupationLevel but low Age and Tenure

C. The observation is an outlier that must be removed

D. The observation has high values on all five variables

Answer: D

A point in the upper-right quadrant has high positive scores on both PC1 (horizontal) and PC2 (vertical). Since Income, Education, and OccupationLevel load heavily on PC1 (pointing right) and Age and Tenure load heavily on PC2 (pointing up), this observation has high values on all five variables. Projecting the point onto each arrow confirms high values in every direction.

Question 12

Which statement best captures the fundamental conceptual difference between PCA and Exploratory Factor Analysis?

A. PCA requires normally distributed data; EFA does not

B. PCA always uses orthogonal rotation; EFA always uses oblique rotation

C. PCA creates components as linear combinations of observed variables; EFA posits that latent factors cause the observed variables

D. PCA can only be used with two variables; EFA works with any number

Question 12

Which statement best captures the fundamental conceptual difference between PCA and Exploratory Factor Analysis?

A. PCA requires normally distributed data; EFA does not

B. PCA always uses orthogonal rotation; EFA always uses oblique rotation

C. PCA creates components as linear combinations of observed variables; EFA posits that latent factors cause the observed variables

D. PCA can only be used with two variables; EFA works with any number

Answer: C

The causal direction is the key distinction. In PCA, components are defined as weighted sums of the observed variables (observed variables produce components). In EFA, the model assumes that unobserved latent factors generate the observed variables, with an error term capturing unique variance. This difference is reflected in their equations: PCA uses $PC = XW$, while EFA uses $X = \Lambda F + e$.

Question 13

A psychologist developing a personality questionnaire suspects that responses are driven by a few underlying traits. She wants to identify these traits and allow them to be correlated. Which approach is most appropriate?

A. EFA with Promax rotation

B. PCA with Varimax rotation

C. PCA with no rotation

D. Simple correlation analysis between all item pairs

Question 13

A psychologist developing a personality questionnaire suspects that responses are driven by a few underlying traits. She wants to identify these traits and allow them to be correlated. Which approach is most appropriate?

A. EFA with Promax rotation

B. PCA with Varimax rotation

C. PCA with no rotation

D. Simple correlation analysis between all item pairs

Answer: A

EFA is the correct framework because the psychologist hypothesizes latent constructs (personality traits) underlying the observed items. Promax rotation is appropriate because personality traits are typically correlated in practice (e.g., extraversion and openness often co-occur). The lesson explicitly recommends Promax when constructs are expected to correlate.

Question 14

In the EFA model $X_i = \lambda_{i1} F_1 + \lambda_{i2} F_2 + e_i$, what does the error term e_i represent?

A. The mean of variable X_i

B. The portion of X_i not explained by the common factors, including measurement error and specific variance

C. The variance explained by the first factor only

D. The correlation between F_1 and F_2

Question 14

In the EFA model $X_i = \lambda_{i1} F_1 + \lambda_{i2} F_2 + e_i$, what does the error term e_i represent?

A. The mean of variable X_i

B. The portion of X_i not explained by the common factors, including measurement error and specific variance

C. The variance explained by the first factor only

D. The correlation between F_1 and F_2

Answer: B

The error term e_i captures all variance in variable X_i that is not accounted for by the common factors. This includes both random measurement error and any variance unique to that specific variable. The lesson states that e_i is assumed to be uncorrelated with the factors F and typically uncorrelated across variables.

Question 15

A variable has factor loadings of 0.6 on Factor 1 and 0.5 on Factor 2 in a two-factor EFA solution. What is its communality h^2?

A. 0.11

B. 0.36

C. 1.10

D. 0.61

Question 15

A variable has factor loadings of 0.6 on Factor 1 and 0.5 on Factor 2 in a two-factor EFA solution. What is its communality h^2?

A. 0.11

B. 0.36

C. 1.10

D. 0.61

Answer: D

Communality is the sum of squared loadings across all factors: $h^2 = 0.6^2 + 0.5^2 = 0.36 + 0.25 = 0.61$. This means 61% of the variable's variance is explained by the two common factors. The remaining 39% is uniqueness ($u^2 = 1 - 0.61 = 0.39$), which includes measurement error and variable-specific variance.

Question 16

A researcher runs a two-factor EFA and finds that variable X5 has a communality of $h^2 = 0.15$. What is the most reasonable interpretation and action?

A. X5 is poorly explained by the factor solution; consider removing it or investigating why

B. X5 is well explained by the factors and should receive the highest weight

C. X5 shares 85% of its variance with the factors and is the most important variable

D. The two-factor solution has too many factors and should be reduced to one

Question 16

A researcher runs a two-factor EFA and finds that variable X5 has a communality of $h^2 = 0.15$. What is the most reasonable interpretation and action?

A. X5 is poorly explained by the factor solution; consider removing it or investigating why

B. X5 is well explained by the factors and should receive the highest weight

C. X5 shares 85% of its variance with the factors and is the most important variable

D. The two-factor solution has too many factors and should be reduced to one

Answer: A

A communality of 0.15 means only 15% of X5's variance is explained by the common factors, while 85% is unique variance. This indicates that X5 does not fit well within the factor structure. The standard practice is to consider removing such items or examining whether they belong to a different construct not captured by the current model.

Question 17

A researcher performs EFA and obtains an unrotated loading matrix where most items load moderately on both Factor 1 and Factor 2. After applying Varimax rotation, the loadings become high on one factor and near zero on the other. What property of the rotated solution has been achieved?

A. Maximum likelihood estimation

B. Increased total variance explained

C. Simple structure

D. Oblique factor correlations

Question 17

A researcher performs EFA and obtains an unrotated loading matrix where most items load moderately on both Factor 1 and Factor 2. After applying Varimax rotation, the loadings become high on one factor and near zero on the other. What property of the rotated solution has been achieved?

A. Maximum likelihood estimation

B. Increased total variance explained

C. Simple structure

D. Oblique factor correlations

Answer: C

Simple structure means each variable loads highly on one factor and has near-zero loadings on the others. Varimax rotation achieves this by maximizing the variance of squared loadings within each factor, pushing high loadings higher and low loadings lower. Importantly, rotation does not change the total variance explained or the model fit; it only redistributes variance among factors to improve interpretability.

Question 18

After an EFA with Promax rotation, the factor correlation matrix Φ shows a correlation of 0.65 between Factor 1 and Factor 2. What does this imply, and what would happen if Varimax had been used instead?

A. The factors are independent; Varimax would produce the same result

B. The factor solution is invalid because correlations above 0.5 are not allowed

C. Promax has inflated the correlation artificially; the true correlation is always zero

D. The factors overlap substantially; Varimax would have forced this correlation to zero, potentially distorting the loading pattern

Question 18

After an EFA with Promax rotation, the factor correlation matrix Φ shows a correlation of 0.65 between Factor 1 and Factor 2. What does this imply, and what would happen if Varimax had been used instead?

A. The factors are independent; Varimax would produce the same result

B. The factor solution is invalid because correlations above 0.5 are not allowed

C. Promax has inflated the correlation artificially; the true correlation is always zero

D. The factors overlap substantially; Varimax would have forced this correlation to zero, potentially distorting the loading pattern

Answer: D

A factor correlation of 0.65 indicates substantial overlap between the two constructs. Varimax, being an orthogonal rotation, constrains all factor correlations to exactly zero, which would misrepresent the true relationship and potentially distort the loading pattern. The lesson teaches that Promax is more realistic when constructs genuinely correlate, as in personality or cognitive ability research.

Question 19

A PCA summary in R shows: PC1 standard deviation = 1.81, PC2 = 1.20, PC3 = 0.98, across 7 standardized variables. What is the cumulative variance explained by PC1 and PC2 combined?

A. 43.0%
B. 53.4%
C. 67.5%
D. 73.8%

Question 19

A PCA summary in R shows: PC1 standard deviation = 1.81, PC2 = 1.20, PC3 = 0.98, across 7 standardized variables. What is the cumulative variance explained by PC1 and PC2 combined?

A. 43.0%

B. 53.4%

C. 67.5%

D. 73.8%

Answer: C

Eigenvalues are the squared standard deviations: $\lambda_1 = 1.81^2 = 3.2761$, $\lambda_2 = 1.20^2 = 1.44$. With 7 standardized variables, total variance $= 7$. Cumulative proportion for PC1 + PC2 $= (3.2761 + 1.44)/7 = 4.7161/7 \approx 0.674$, or approximately 67.5%. This means two components capture about two-thirds of the information in seven original variables.

Question 20

When reporting an EFA in a research paper, which combination of information is essential to include?

A. Only the eigenvalues and the scree plot

B. The rotated loading matrix, communalities, percentage of variance explained, rotation method used, and factor correlation matrix (if oblique)

C. Only the number of factors and their names

D. The raw data matrix and all unrotated loadings

Question 20

When reporting an EFA in a research paper, which combination of information is essential to include?

A. Only the eigenvalues and the scree plot

B. The rotated loading matrix, communalities, percentage of variance explained, rotation method used, and factor correlation matrix (if oblique)

C. Only the number of factors and their names

D. The raw data matrix and all unrotated loadings

Answer: B

A complete EFA report must include the rotated factor loading matrix (showing which items belong to which factor), communalities (showing how well each item is explained), total and per-factor variance explained, the rotation method (Varimax, Promax, etc.), and the factor correlation matrix Φ when oblique rotation is used. The lesson explicitly states that Φ must be reported for Promax solutions.