

Pontificia Universidad Javeriana
Departamento de Matemáticas
Análisis Multivariado. 3er parcial

Las autoridades de un banco desean construir un modelo con el que puedan identificar si los billetes son falsos o no utilizando algunas características físicas de los mismos. La idea es encontrar el tal modelo y con el construir un algoritmo para programar un escaner que tome las medidas e identifique si el billete pudiera ser considerado como falso o no. Construya el modelo utilizando el análisis discriminante, valor su capacidad de discriminación y utilícelo para describir las características de cada uno de los tipos de billetes

Las medidas consideradas fueron:

longitud: longitud del billete

long_izq: ancho de la parte izquierda del billete

long_der: ancho de la parte derecha del billete

margen_inf: margen de la parte inferior del billete

margen_sup: margen de la parte superior del billete

diag: longitud diagonal del billete

tipo: falso, real

Para responder las preguntas debe utilizar el archivo pregunta 3 que contiene los resultados del procesamiento computacional.

Pregunta 1 (0.4 puntos)

Diga si la afirmación siguiente es verdadera o falsa.

En la función discriminante de Fisher la variable más importante para la discriminación de los billetes es la variable **longitud**

☐ Verdadero

☒ Falso

Pregunta 2 (0.4 puntos)

Diga si la afirmación siguiente es verdadera o falsa.

Utilizando el modelo de discriminación construido, el porcentaje de clasificaciones de los billetes que son correctas es del 99.5%

☒ Verdadero

☐ Falso

Pregunta 3 (0.7 puntos)

Los billetes falsos se caracterizan por:

- ☐ ser más largos
- ☒ tener márgenes (inferior y superior) mayores, menos longitud diagonal, ser más anchos en la parte derecha y menos anchos en la parte izquierda
- ☐ tener márgenes (inferior y superior) menores, mayor longitud diagonal, ser menos anchos en la parte derecha y más anchos en la parte izquierda
- ☐ tener mayores dimensiones en general
- ☐ ninguna de las anteriores

Pregunta 4 (0.5 puntos)

Un billete que tenga:

longitud=214.6

long_izq= 130

long_der=130.5

margen_inf= 8

margen_sup = 10.5

diag=140

Si las probabilidades a posteriori de pertenecer a cada grupo son:

falso	real
0.0435	0.9564

clasificaría como

- ☒ real
- ☐ falso

Un grupo de investigación en recursos humanos desean determinar los rasgos (factores) que distinguen a los candidatos a puestos de trabajo, para lograr su objetivo miden 12 variables y realizan un análisis factorial para identificar los factores que caracterizan el comportamiento conjunto de las diferentes variables analizadas

Variables

acad: nivel académico
 apariencia: apariencia física
 comunicación: comunicación entre personas
 compatibilidad: compatibilidad con los intereses de la compañía
 experiencia: experiencia laboral
 calidad: calidad en su desempeño laboral
 escritura: habilidad al escribir textos
 sociabilidad: mide lo sociable que es la persona
 organización: mide lo organizada que es la persona
 potencial: posibilidades de mejorar en sus calificaciones
 resume: habilidad al resumir las ideas básicas de un texto
 confianza: autoconfianza de la persona
 Para sus análisis **utilice el archivo pregunta 2** donde aparecen las salidas del procesamiento computacional

Pregunta 5 (0.2 puntos)

Diga si el enunciado siguiente es verdadero o falso.

La prueba de Bartlett indica que hay una relación muy importante entre las variables por lo que es recomendable analizarlas utilizando el análisis factorial

- ☐ Verdadero
☒ Falso

Pregunta 6 (0.2 puntos)

Diga si la afirmación siguiente es verdadera o falsa.

Al aplicar el criterio KMO el coeficiente KMO tiene un valor de 0.8, lo que indica que existe una relación muy débil entre las variables

- ☐ Verdadero
☒ Falso

Pregunta 7 (0.3 puntos)

Diga si la afirmación siguiente es verdadera o falsa.

Al aplicar las medidas de lo adecuado de la muestra (MSA) para identificar las variables con una relación nula o muy débil con las otras. se puede afirmar que la variable **resume** debería ser excluida del análisis por tener un nivel de relación muy bajo.

- ☐ Verdadero
☒ Falso

Pregunta 8 (0.2 puntos)

Si se consideran 6 factores para el análisis , el porcentaje de la varianza total que se logra explicar con ellos es

- ☐ 7%
- ☐ 100%
- ☐ Ninguno de los anteriores
- ☒ 86%
- ☐ 90%

Pregunta 9 (0.2 puntos)

El porcentaje de la varianza total explicado por el factor denominado ML5 es:

- ☐ 22%
- ☐ Ninguno de los anteriores
- ☒ 15%
- ☐ 19%
- ☐ 14%

Pregunta 10 (0.2 puntos)

Diga si la afirmación siguiente es verdadera o falsa.

El factor que explica el mayor porcentaje de la varianza total es el denominado ML5

- ☐ Verdadero
- ☒ Falso

Pregunta 11 (0.2 puntos)

La correlación entre la **variable acad** y el **factor denominado ML1** es

- ☐ 0.141
- ☐ Ninguno de los anteriores
- ☒ 0.17
- ☐ 0.86

Pregunta 12 (0.3 puntos)

Diga cuál de las afirmaciones siguientes se corresponde con la interpretación del factor denominado ML1

- ☒ Habilidades al escribir
- ☐ Características de la personalidad
- ☐ Ninguna de las anteriores
- ☐ Capacidad técnica y potencial de desarrollo

Pregunta 13 (0.2 puntos)

La varianza de la variable comunicación explicada por los 6 factores es:

- ☒ 0.79
- ☐ 0.210
- ☐ ninguna de las anteriores
- ☐ 1.84

En las situaciones siguientes identifique cuál sería el procedimiento de análisis más adecuado:

Pregunta 14 (0.2 puntos)

Se estudian diferentes variables que caracterizan la personalidad (como por ejemplo sociabilidad, espontaneidad, curiosidad, nerviosismo, agresividad etc.) con el fin de identificar rasgos distintivos de la misma

- ☐ Análisis de Conglomerados
- ☐ Análisis Discriminante
- ☒ Análisis Factorial
- ☐ Correlaciones canónicas
- ☐ Ninguna de las anteriores

Pregunta 15 (0.1 puntos)

Una compañía de computación está trabajando en la elaboración de algoritmos para comprimir imágenes. En este caso cada imagen es tratada como un conjunto de datos multivariado donde cada pixel representa una variable y les interesa reducir esta gran cantidad de variables en un conjunto más pequeño pero que mantenga la mayor información posible del conjunto de variables original

- ☐ Análisis Factorial
- ☒ Ninguna de las anteriores
- ☐ Correlaciones canónicas
- ☐ Análisis Discriminante
- ☐ Análisis de Conglomerados

Pregunta 16 (0.2 puntos)

Una compañía dedicada al transporte de cargas internacionales agrupa a sus empleados, según sus perfiles de trabajo, en tres grupos; (1) atención al cliente (2) mecánicos (3) despachadores. Para caracterizar cada uno de los grupos y determinar si la pertenencia a cada uno de ellos se corresponde con diferentes tipos de personalidad de los empleados, recoge información (diferentes variables) sobre actividad extra-laboral, sociabilidad y conservadurismo

- ☐ Correlaciones canónicas
- ☐ Análisis Factorial
- ☒ Análisis Discriminante
- ☐ Análisis de Conglomerados
- ☐ Ninguna de las anteriores

Pregunta 17 (0.1 puntos)

Los directivos de una universidad acostumbran a clasificar a sus estudiantes en tres grupos de habilidades educativas. Con el fin de crear un modelo que les permita clasificar a los estudiantes en el momento de matricular. Toman una muestra de 180 estudiantes (60 en cada uno de los grupos) y les aplican dos pruebas, las que brindan información (con eso construyen dos variables), respectivamente, sobre rendimiento académico y motivación

- ☐ Correlaciones canónicas
- ☐ Análisis de conglomerados
- ☐ Ninguna de las anteriores
- ☒ Análisis Discriminante
- ☐ Análisis Factorial

Pregunta 18 (0.2 puntos)

Un banco recoge información sobre las personas que solicitan créditos (como por ejemplo ingresos, educación, edad, deudas en tarjetas de crédito, etc) con el fin de encontrar un número mínimo de índices que le permita disponer de la mayoría de la información sobre los aplicantes con el número mínimo de índices posibles.

- ☐ Análisis Factorial
- ☐ Correlaciones Canónicas
- ☐ Análisis Discriminante
- ☒ Ninguna de las anteriores
- ☐ Análisis de Conglomerados

Pregunta 19 (0.2 puntos)

Se estudian diferentes variables que caracterizan los tipos de costos de producción de una importante compañía industrial (como por ejemplo costo de materiales, salarios, costo de equipos, costos fijos, etc) con el fin de identificar características distintivas de los costos.

- ☐ Correlaciones canónicas
- ☒ Análisis Factorial
- ☐ Análisis de conglomerados
- ☐ Ninguna de las anteriores
- ☐ Análisis Discriminante