



UNIVERSIDAD NACIONAL DE INGENIERIA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA - FIEECS

ESTADÍSTICA MULTIVARIADA

Dr. Luis Huamanchumo de la Cuba

Seminario Pre Final

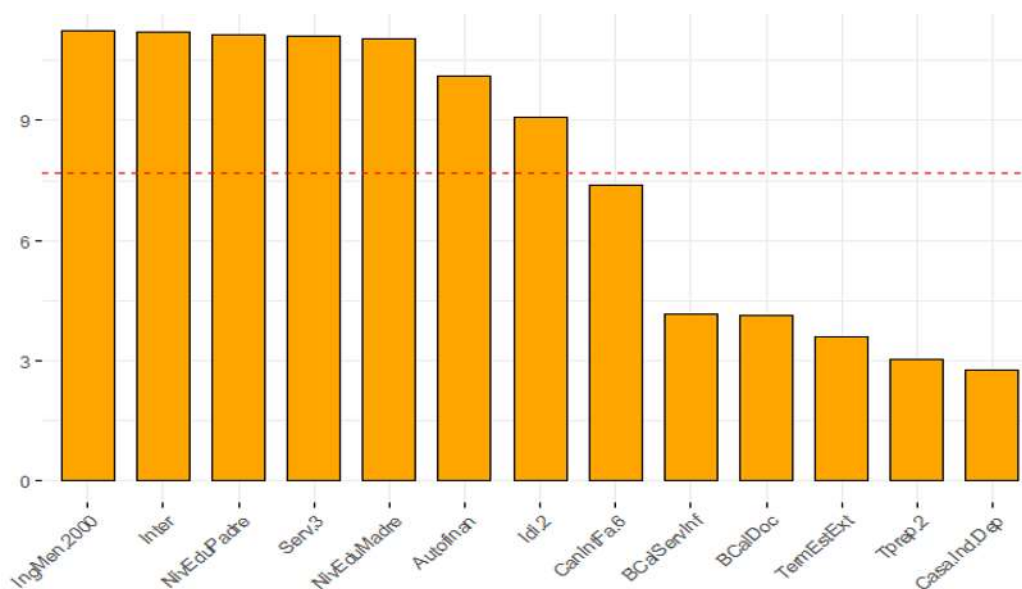
1. ¿En qué consiste la “ambigüedad inherente” en el análisis factorial?
2. Es verdadera, falsa o incierta la siguiente afirmación: “una rotación rígida reproduce la matriz de covarianzas”. Justifique su respuesta.
3. Un modelo factorial puede calificarse en la medida en que las cargas factoriales reproducen de manera efectiva los datos originales.
4. Se presenta los resultados de un estudio en Universidades peruanas de la provincia de Lima y Provincia Constitucional del Callao. Explique con sustento teórico los resultados presentados. Luego, a partir de los resultados sustentados plantee una hipótesis.

INDICADORES	VARIABLE	DESCRIPCIÓN
Idi>2 :	Var1	Saben al menos 2 idiomas aparte del español
Serv>3 :	Var2	Cuentan con al menos 3 de los 4 servicios básicos
Inter :	Var3	Cuentan con servicio de Internet
Casa Ind/Dep :	Var4	Estudiante cuenta con Casa Independiente o Departamento en Edificio
Autofinan :	Var5	Porcentaje de alumnos que autofinancian su carrera
BCalServInf :	Var6	Buena calidad de servicios informaticos
TermEstExt :	Var7	Piensa terminar sus estudios universitarios en el extranjero
BCalDoc :	Var8	Buena calidad de plana docente
Tprep>2 :	Var9	Tiempo de preparación mayor o igual a 2 años
IngMen>2000 :	Var10	Ingreso mensual del hogar > 2000
NivEduPadre:	Var11	Nivel de educación del padre superior a universitaria completa
NivEduMadre:	Var12	Nivel de educación del madre superior a universitaria completa
CanIntFa≥6:	Var13	Cantidad de integrantes mayor a 6 por familia

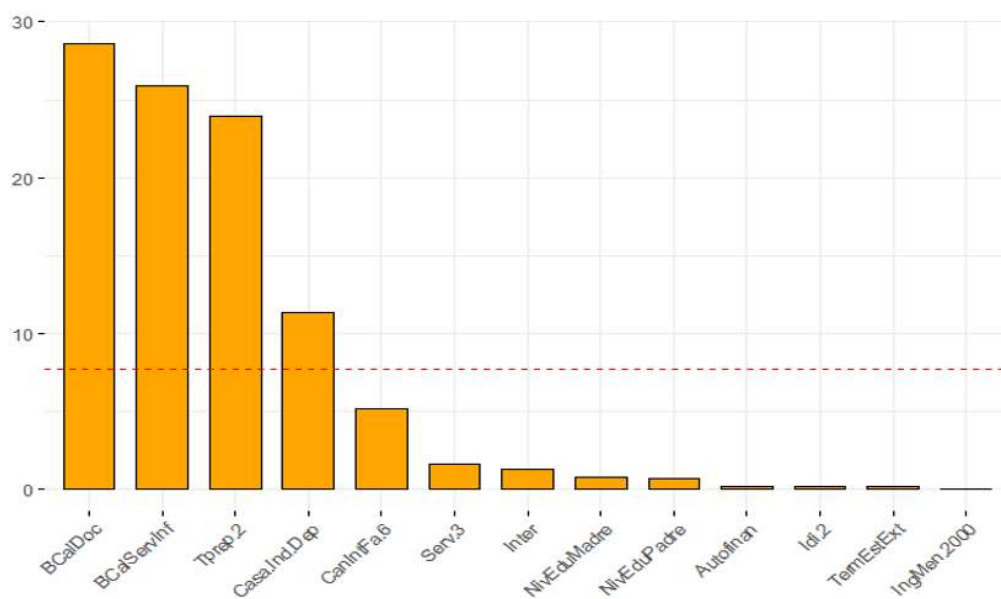
Figura 1.- Indicadores de universidades de Lima y Callao

OBJ	UNIVERSIDADES	
1	PUCP:	PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
2	UARM:	UNIVERSIDAD ANTONIO RUÍZ DE MONTTOYA
3	UADP:	UNIVERSIDAD AUTÓNOMA DEL PERÚ
4	UCSS:	UNIVERSIDAD CATÓLICA SEDES SAPIENTIAE
5	UCS:	UNIVERSIDAD CIENTÍFICA DEL SUR
6	UESAN:	UNIVERSIDAD DE ADMINISTRACIÓN DE NEGOCIOS ESAN
7	UCH:	UNIVERSIDAD DE CIENCIAS Y HUMANIDADES
8	ULima:	UNIVERSIDAD DE LIMA
9	USMP:	UNIVERSIDAD DE SAN MARTÍN DE PORRES
10	UP:	UNIVERSIDAD DEL PACÍFICO
11	UNIFE:	UNIVERSIDAD FEMENINA DEL SAGRADO CORAZÓN
12	UJBM:	UNIVERSIDAD JAIME BAUSATE Y MEZA
13	ULCB:	UNIVERSIDAD LE CORDON BLEU
14	UNALM:	UNIVERSIDAD NACIONAL AGRARIA LA MOLINA
15	CANTUTA:	UNIVERSIDAD NACIONAL DE EDUCACIÓN ENRIQUE GUZMAN Y VALLE
16	UNI:	UNIVERSIDAD NACIONAL DE INGENIERÍA
17	UNAC:	UNIVERSIDAD NACIONAL DEL CALLAO
18	UNFV:	UNIVERSIDAD NACIONAL FEDERICO VILLARREAL
19	UNMSM:	UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
20	UPCH:	UNIVERSIDAD PERUANA CAYETANO HEREDIA
21	UPC:	UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS
22	UPeU:	UNIVERSIDAD PERUANA UNIÓN
23	UPNW:	UNIVERSIDAD PRIVADA NORBERT WIENER
24	USIL:	UNIVERSIDAD PRIVADA SAN IGNACIO DE LOYOLA
25	URP:	UNIVERSIDAD RICARDO PALMA
26	UTP:	UNIVERSIDAD TECNOLÓGICA DEL PERÚ

Figura 2.- Universidades de Lima y Callao



(a)



(b)

Gráfico 01.- Calidad de representación porcentual de las variables con a) la primera componente, b) segunda componente

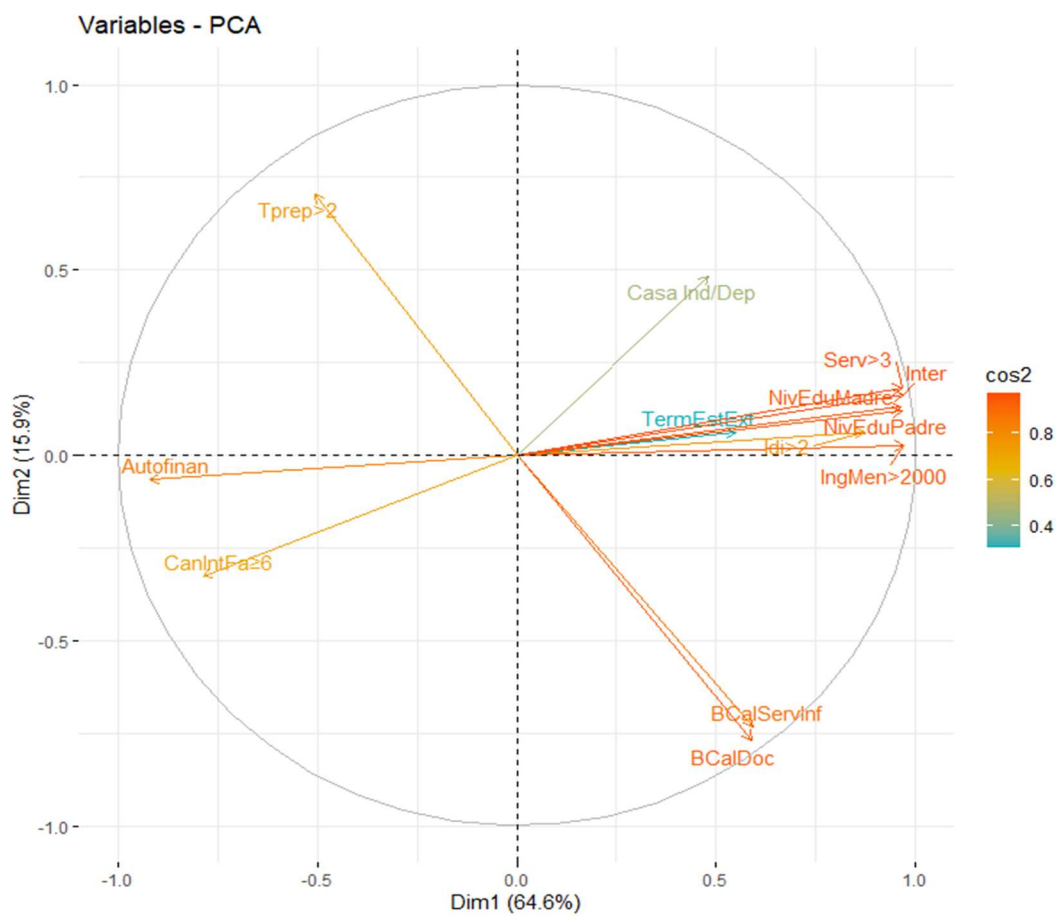


Gráfico 02.- Gráfico de Variables en el mapa 80.5%

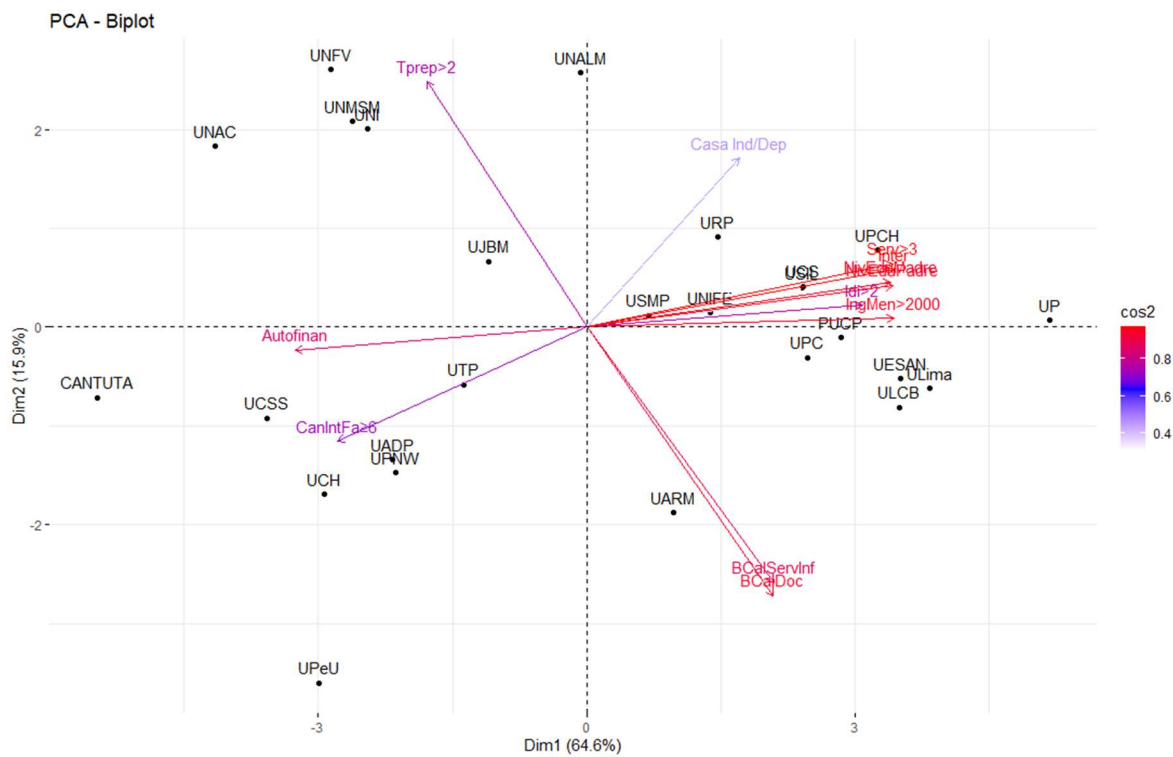


Gráfico 03.- Biplot de Análisis de Componentes Principales

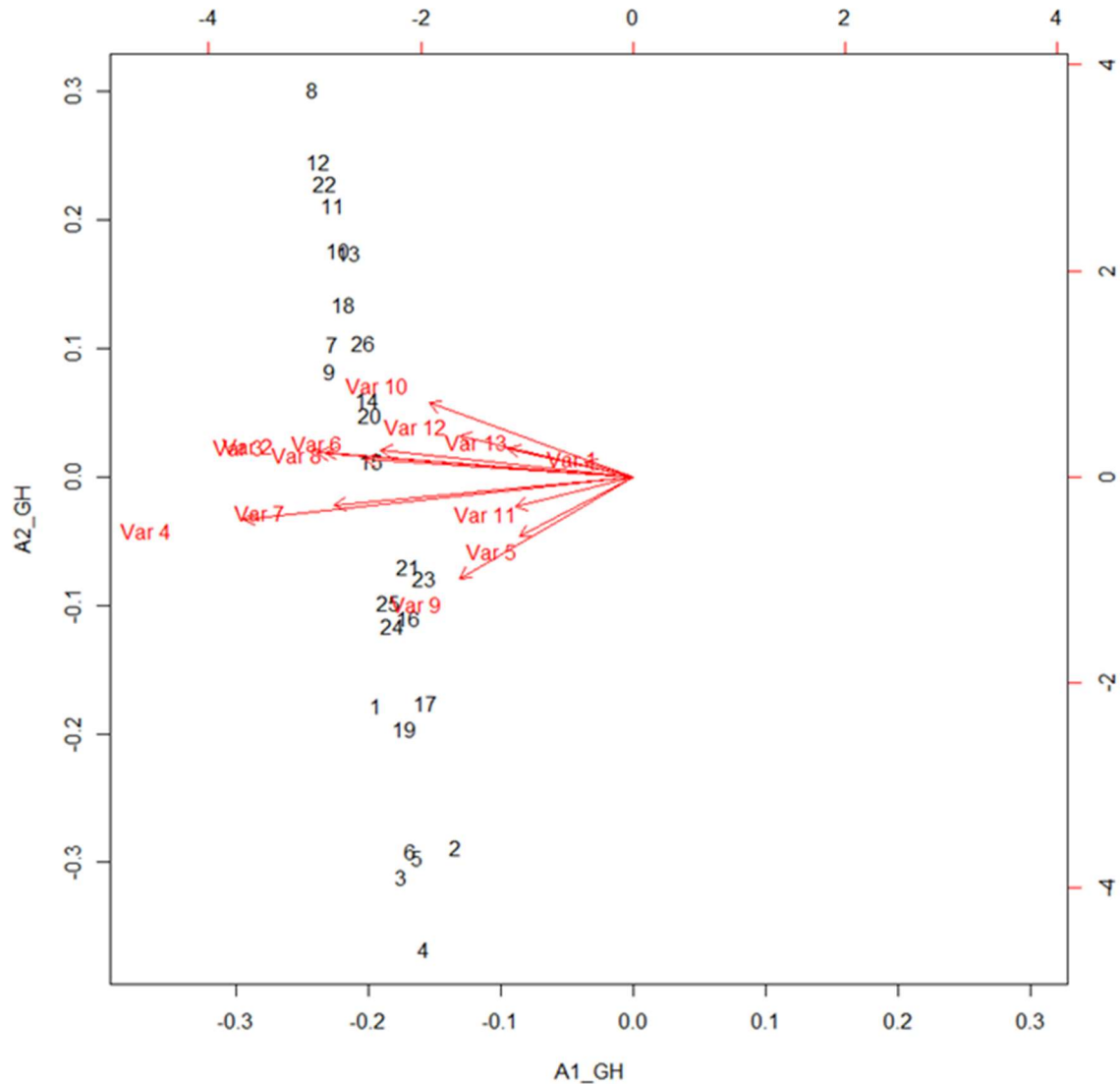


Gráfico 04.- GH Biplot. Variables (figura 1) y universidades (figura 2)

5. Explique cada uno de los casos presentados para el análisis biplot. Trate de llegar a un grado de formalidad en su explicación. Puede complementar tomando como ejemplo la exposición en clase.

$$X=AB'$$

Representación gráfica de **X**
en dos dimensiones

$$X=U*L^{c/2}L^{(1-c)/2}U'$$

$$0 \leq c \leq 1$$

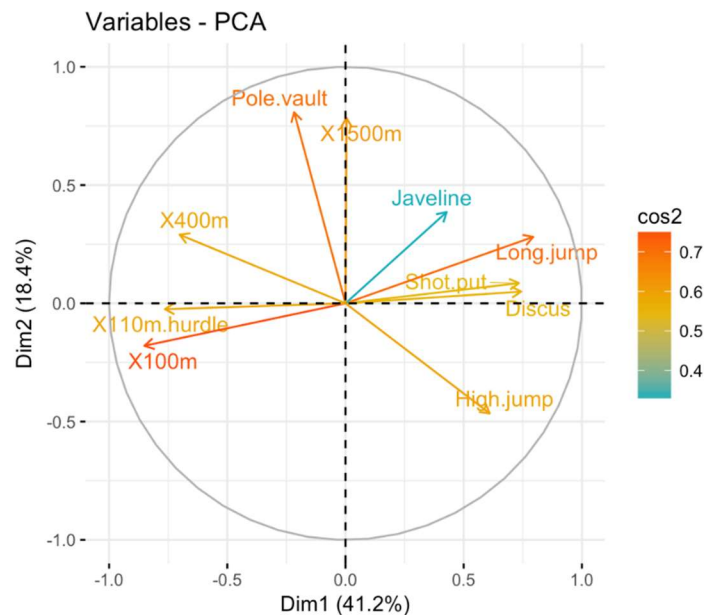
$$A=U*L^{c/2} \text{ y } B'=L^{(1-c)/2}U'$$

C=0.5 SQ biplot $A=U*L^{1/4}$ y $B'=L^{1/4}U'$ Para diagnóstico

C=0 GH biplot $A=U^*$ y $B'=L^{1/2}U'$ Preserva la métrica de columna

C=1 JK biplot $A=V^*=Z$ y $B'=U'=Y^*$ Preserva la métrica de filas

6. Los elementos de la matriz $D = d_{rs}$ son distancias euclideas, si y solo si, la matriz de doble centralidad $B = HAH$ es definida no negativa. Demuestre la suficiencia.
7. Demuestre que el GH biplot preserva la métrica de columnas.
8. En el gráfico adjunto explique cómo se obtiene la calidad de representación de las variables.



9. La siguiente afirmación es verdadera o falsa, sustente rigurosamente su respuesta.
“En el análisis biplot la distancia de filas puede verse como una distancia ponderada de las observaciones originales”.

10. En el modelo factorial cuál es la relación entre los puntajes de los factores cuando se aplica una rotación de cargas.
11. En el análisis factorial con el método de estimación de componentes principales deduzca de manera detallada la contribución a la varianza total muestral del j-ésimo factor.
12. La matriz de doble centralidad definida no negativa es suficiente para que la matriz de distancias esté definida por distancias euclideas.
13. En el modelo factorial, determinar el estimador mínimos cuadrados ponderados de los factores comunes estableciendo los supuestos que crea conveniente.
14. Suponga que para el modelo:

$$\mathbf{X}(p \times 1) - \boldsymbol{\mu}(p \times 1) = \mathbf{L}(p \times m) \mathbf{F}(m \times 1) + \boldsymbol{\varepsilon}(p \times 1)$$
 el vector de medias $\boldsymbol{\mu}$, las cargas factoriales \mathbf{L} y la varianza específica $\boldsymbol{\psi}$ se estiman por Máximo Verosimilitud. Muestre que cuando a las cargas factoriales se les aplica una rotación rígida los puntajes factoriales en ambos espacios obtenidos por regresión se relacionan mediante una transformación lineal.
15. Comente por qué la suma de cuadrados de todos los elementos de la matriz residual ($p \times p$) no debe ser mayor que la suma de los cuadrados de los $p-m$ últimos valores característicos cuando se trata de determinar la cantidad de factores m en el modelo factorial.
16. Un modelo factorial puede calificarse en la medida en que las cargas factoriales reproducen de manera efectiva la matriz de covarianzas.
17. "Identificar clusters sobre la base de la minimización de la distancia euclídea es equivalente a maximizar las correlaciones lineales".