



UNIVERSIDAD DE PUERTO RICO
Recinto Rio piedras



Iniciativas de Investigación y Actividad Creativa Subgraduada (iINAS Título V Subgraduado)

Investigación cuantitativa usando



Parte II

Presentado por: José Carlos Vega Vilca

¿CÓMO INSTALAR ?

<http://cran.r-project.org/>

- 1) Click en **Download R for Windows**
- 2) Click en **base**
- 3) Click en **Download R 3.1.1 for Windows**
- 4) Click en **Run**

DESKTOP: Dar doble click en



DATOS REALES:

Egresados FAE-2009

IGS: Índice General de Solicitud

PromGrad: Promedio de Graduación

AñosEstu: Número de años para culminar estudios

MECU: Puntuación obtenida en los cursos MECU-3031 y MECU-3032

ESCUELA: Tipo de escuela. 0 = Pública, 1 = Privada

GENERO: Género del estudiante. 0 = Hombre, 1 = Mujer

EGRESADOS FAE-2009

Desktop de su computadora

Click en Folder “Taller iINAS”

Click en DatosTaller, EXCEL

IMPORTAR DATOS: DESDE EXCEL HACIA

1) En EXCEL, seleccionar los datos, COPY

Seleccionar los datos de la variables “IGS”, “PromGrad”, “AñosEstu”,
“MECU”, “Escuela”, “Genero”.

2) En escribir:

```
datos=read.table("clipboard",header=T)
```

```
head(datos)
```

```
names(datos)
```

```
IGS
```

```
attach(datos)
```

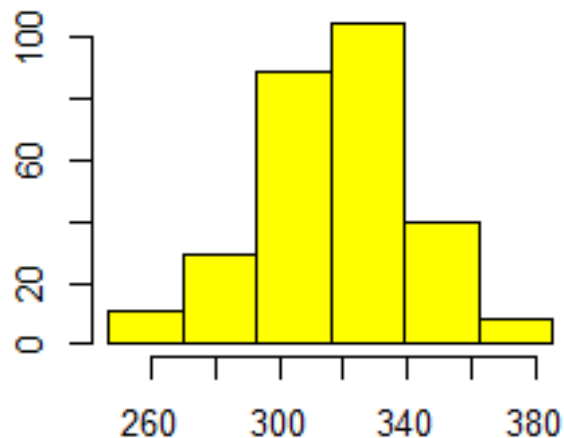
Normalidad de los datos: IGS

Ho: Los datos SI tienen distribución Normal

H1: Los datos NO tienen distribución Normal

MÉTODO GRÁFICO:

Histograma de frecuencias



MÉTODO ANALÍTICO:

Prueba de Shapiro-Wilk

```
> shapiro.test(IGS)
```

Shapiro-Wilk normality test

data: IGS

W = 0.993, **p-value = 0.2079**

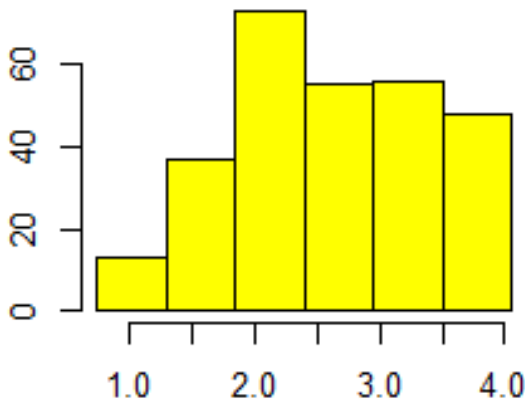
Normalidad de los datos: MECU

Ho: Los datos SI tienen distribución Normal

H1: Los datos NO tienen distribución Normal

MÉTODO GRÁFICO:

Histograma de frecuencias



MÉTODO ANALÍTICO:

Prueba de Shapiro-Wilk

```
> shapiro.test(MECU)
```

Shapiro-Wilk normality test

```
data: MECU
```

```
W = 0.9621, p-value = 9.655e-07
```

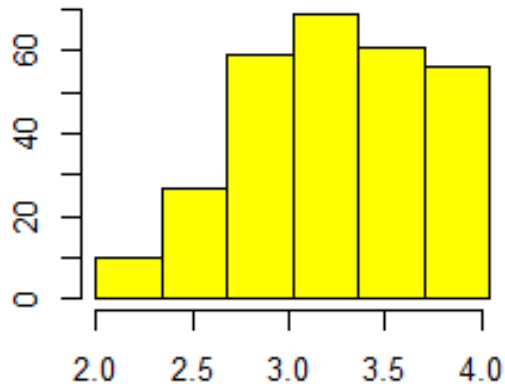
Normalidad de los datos: PromGrad

Ho: Los datos SI tienen distribución Normal

H1: Los datos NO tienen distribución Normal

MÉTODO GRÁFICO:

Histograma de frecuencias



MÉTODO ANALÍTICO:

Prueba de Shapiro-Wilk

```
> shapiro.test(PromGrad)
```

Shapiro-Wilk normality test

```
data: PromGrad
```

```
W = 0.9664, p-value = 3.808e-06
```


¿VARIANZAS IGUALES? (en dos grupos)

Datos: IGS

$H_0 : \sigma^2_{pública} = \sigma^2_{privada}$

$H_1 : \sigma^2_{pública} \neq \sigma^2_{privada}$

p-value = 0.2858

$H_0 : \sigma^2_{hombres} = \sigma^2_{mujeres}$

$H_1 : \sigma^2_{hombres} \neq \sigma^2_{mujeres}$

p-value = 0.9769

¿VARIANZAS IGUALES? (en dos grupos)

Datos: MECU

$H_0 : \sigma_{\text{pública}}^2 = \sigma_{\text{privada}}^2$

$H_1 : \sigma_{\text{pública}}^2 \neq \sigma_{\text{privada}}^2$

p-value = 0.08292

$H_0 : \sigma_{\text{hombres}}^2 = \sigma_{\text{mujeres}}^2$

$H_1 : \sigma_{\text{hombres}}^2 \neq \sigma_{\text{mujeres}}^2$

p-value = 0.4714

¿VARIANZAS IGUALES? (en dos grupos)

Datos: Promedio de Graduación

$H_0 : \sigma^2_{pública} = \sigma^2_{privada}$

$H_1 : \sigma^2_{pública} \neq \sigma^2_{privada}$

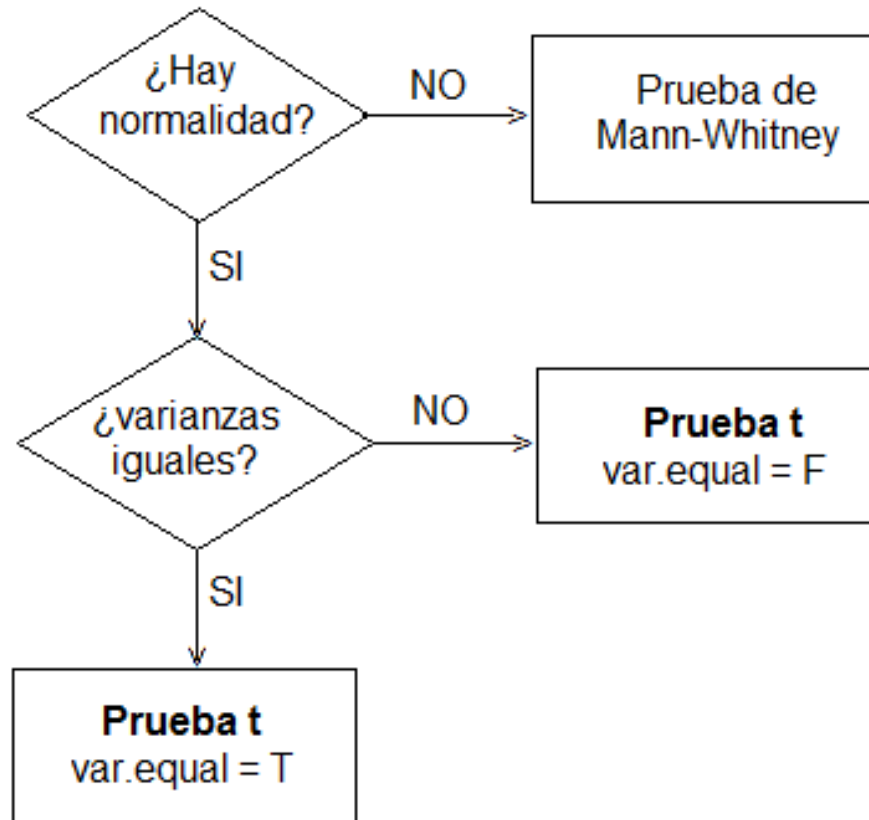
p-value = 0.4294

$H_0 : \sigma^2_{hombres} = \sigma^2_{mujeres}$

$H_1 : \sigma^2_{hombres} \neq \sigma^2_{mujeres}$

p-value = 0.07662

Prueba de Diferencia de Medias



¿EXISTE DIFERENCIA DE MEDIAS?

Datos: *IGS* { Normalidad
iguales

Varianzas

Prueba T

$H_0 : \mu_{\downarrow pública} \uparrow = \mu_{\downarrow privada} \uparrow$

$H_1 : \mu_{\downarrow pública} \uparrow \neq \mu_{\downarrow privada} \uparrow$

p-value = 0.6074

$H_0 : \mu_{\downarrow hombres} \uparrow = \mu_{\downarrow mujeres} \uparrow$

$H_1 : \mu_{\downarrow hombres} \uparrow \neq \mu_{\downarrow mujeres} \uparrow$

p-value = 0.008481

	Media	Mediana
Hombres	311.7	314.5
Mujeres	319.4	318.5

¿EXISTE DIFERENCIA DE MEDIAS?

Datos: *MECU* {■ No hay
Normalidad Varianzas iguales

Prueba de
Mann-Whitney

$H_0 : \mu_{\text{pública}} = \mu_{\text{privada}}$
 $H_1 : \mu_{\text{pública}} \neq \mu_{\text{privada}}$

p-value = 0.7011

$H_0 : \mu_{\text{hombres}} = \mu_{\text{mujeres}}$
 $H_1 : \mu_{\text{hombres}} \neq \mu_{\text{mujeres}}$

p-value = 0.3467

¿EXISTE DIFERENCIA DE MEDIAS?

Datos: *PromGrad* { No hay Normalidad Varianzas iguales

Prueba de Mann-Whitney

$H_0 : \mu_{\text{pública}} = \mu_{\text{privada}}$
 $H_1 : \mu_{\text{pública}} \neq \mu_{\text{privada}}$

$H_0 : \mu_{\text{hombres}} = \mu_{\text{mujeres}}$
 $H_1 : \mu_{\text{hombres}} \neq \mu_{\text{mujeres}}$

p-value = 0.8636

p-value = 0.004768

	Media	Mediana
Hombres	3.15	3.16
Mujeres	3.33	3.33

ANALISIS DE CORRELACION LINEAL

MIDE LA ASOCIACION LINEAL ENTRE DOS VARIABLES

Valor POSITIVO → Relación Directa

Valor NEGATIVO → Relación Inversa

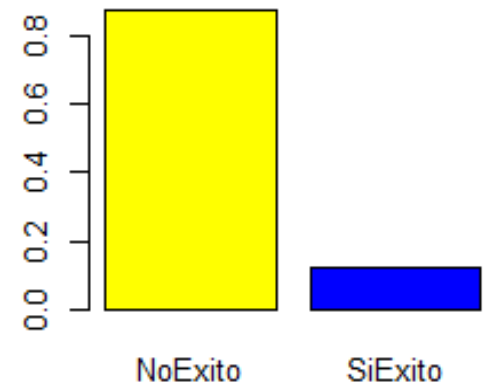
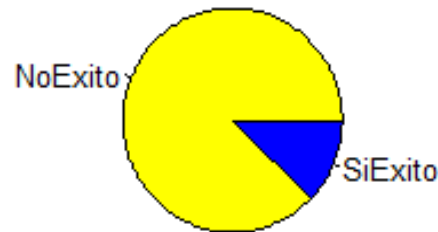
	IGS	PromGrad	AñosEstu	MECU
IGS	1.0000	0.2728	-0.2826	0.3380
PromGrad	0.2728	1.0000	-0.3071	0.3642
AñosEstu	-0.2826	-0.3071	1.0000	-0.2653
MECU	0.3380	0.3642	-0.2653	1.0000

Transformación de variables (1)

EGRESADO DE EXITO: Egresado con Promedio de Graduación de al menos 3.33, que culminó sus estudios en 5 años o menos

exito = (PromGrad >= 3.33 & AñosEstu <= 5) * 1

	Porcentaje
NoExito	87.59
SiExito	12.41



Transformación de variables (2)

Categorización del IGS	1) IGS bajo:	hasta 302
	2) IGS medio:	303 – 333
	3) IGS alto:	334 a más

Categorización del MECU	1) MECU bajo:	hasta 2.00
	2) MECU medio:	2.10 – 3.00
	3) MECU alto:	3.10 a más

```
a1= (IGS<=302) *1
a2= (IGS>=303 & IGS<=333) *2
a3= (IGS>=334) *3

cIGS=a1+a2+a3
```

```
b1= (MECU<=2.00) *1
b2= (MECU>=2.1&MECU<=3.0) *2
b3= (MECU>=3.10) *3

cMECU=b1+b2+b3
```

Tabla de Contingencia

H₀: Las categorías de IGS **NO** están asociadas a las categorías de MECU

H₁: Las categorías de IGS **SI** están asociadas a las categorías de MECU

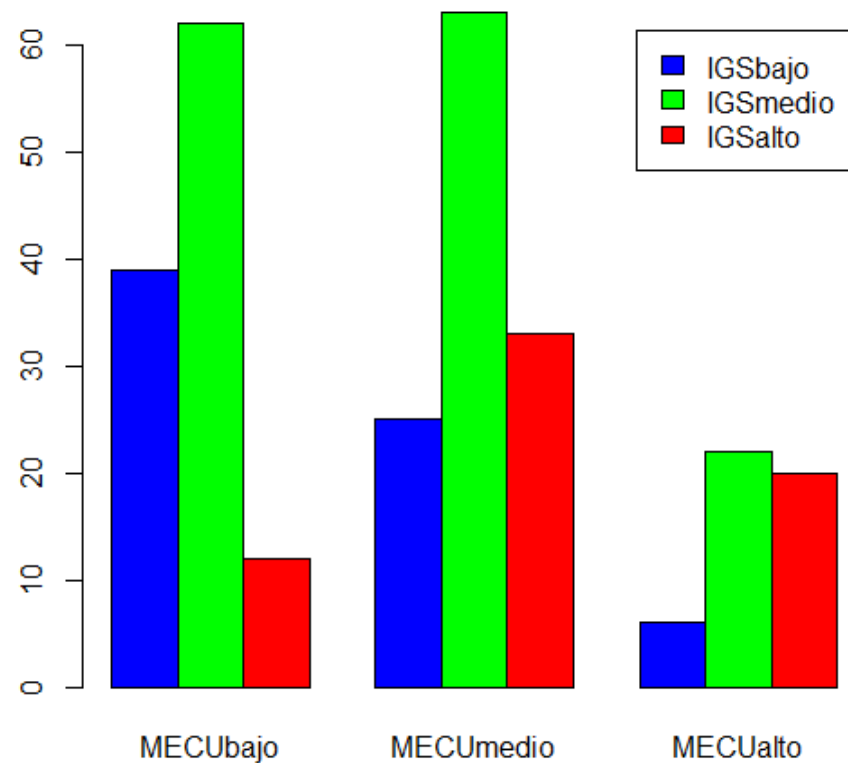
	cMECU		
cIGS	1	2	3
1	39	25	6
2	62	63	22
3	12	33	20

Pearson's Chi-squared test

X-squared = 24.3117

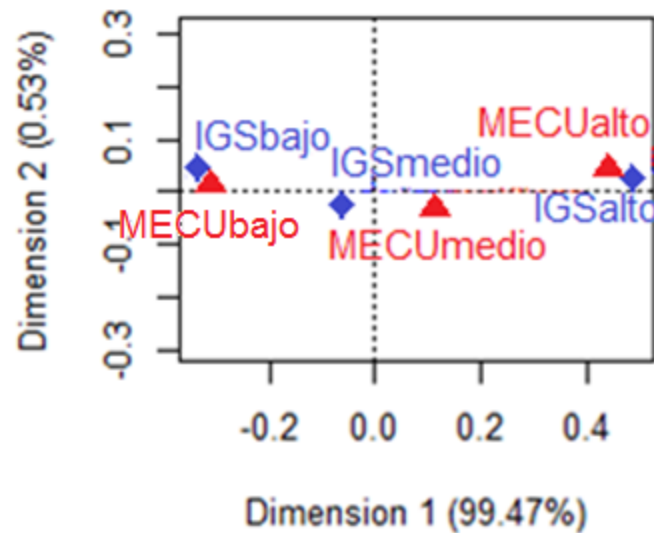
df = 4

p-value = **6.917e-05**



Análisis de Correspondencias Simples

	cMecu		
cIGS	MECUBajo	MECUmedio	MECUalto
IGSbajo	39	25	6
IGSmedio	62	63	22
IGSalto	12	33	20



Regresión Logística (1)

EGRESADO DE EXITO: Egresado con Promedio de Graduación de al menos 3.33, que culminó sus estudios en 5 años o menos

El éxito académico de un egresado parece depender de las siguientes variables: "IGS", "MECU", "Escuela", "Genero",

Regresión Logística (2)

Modelo de Regresión Logística

Call:

```
glm(formula = exito ~ IGS + MECU + Escuela + Genero,  
family = "binomial")
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-18.69692	3.55232	-5.263	1.41e-07	***
IGS	0.04114	0.01072	3.838	0.000124	***
MECU	1.12882	0.30688	3.678	0.000235	***
Escuela	0.47855	0.50346	0.951	0.341854	
Genero	-0.25053	0.43843	-0.571	0.567706	

Regresión Logística (3)

Interpretación de los coeficientes mediante el Odds Ratio (OR)

```
> exp(cbind(OR=coef(mod), confint(mod)))
```

		OR	2.5 %	97.5 %
(Intercept)	7.586350e-09	4.394977e-12	5.275606e-06	
IGS	1.041999e+00	1.021212e+00	1.065255e+00	
MECU	3.092007e+00	1.729995e+00	5.804226e+00	
Escuela	1.613729e+00	6.321045e-01	4.695894e+00	
Genero	7.783871e-01	3.314546e-01	1.877629e+00	

Regresión Logística (4)

CON EL MODELO DE **EXITO ACADEMICO** QUE DEPENDE DE “**IGS**” Y “**MECU**”

Calcular la probabilidad de EXITO ACADEMICO, en los casos:

Estudiante	IGS	MECU
1	300	2.5
2	360	3.5
3	365	4.0

```
casos=data.frame(IGS=c(300,360,365),  
                 MECU=c(2.5,3.5,4.0))  
predict(mod2,casos,type="response")
```


GRACIAS