

Estadística – Atorresi

I.

1- los datos que se reúnen son tomados a 163 mujeres que escuchan radio a las 15 hs.

aparece 2 gráficos de torta o "pictograma según propone botella".

un gráfico indica que los datos fueron tomados a mujeres entre 20 y 30 años para quien no escucha ninguna radio los valores son de 21% y para las que escuchan am 30% y para quienes escuchan pm 41%. 77 mujeres fueron entrevistadas entre 20 y 30 años. En el otro caso las mujeres entre 30 y 60 que respondieron a la encuesta fueron 88, sus porcentajes fueron 21% para quienes no escucharon a ninguna radio 35 para quienes escuchaban am y 45 para quienes escuchaban pm.

a) Determinar población y muestra de individuo

b) calcular tabla de frecuencias relativa, absoluta y porcentual

c) teniendo en cuenta las dos distribuciones, ¿con respecto a que medidas de variabilidad las podrías relacionar y con respecto a que medida de tendencia central? justifique y aplique tales medidas?

3- serán capacitados aquellos operarios que obtengan un puntaje superior a 73 y aquellos que no superen el decil dos.

4- Explique medidas de tendencia central (media, mediana y moda) y su relación con las medidas de tendencia.

RESPUESTA: la moda se puede aplicar a todos los niveles de medición (nominal, ordinal, intervalar y de razón).

la mediana se puede aplicar a los niveles de medición ordinal, intervalar y de razón.

la mediana se puede aplicar a los niveles de medición intervalar o de razón.

5- Explique los niveles de medición y ejemplifique.

Niveles de medición nominal refiere a características diferentes entre sí (color de pelo rubio, morocho, colorado).

ordinal refiere a características de diferente orden.

intervalar su cero es un cero arbitrario (temperatura).

de razón su cero es un cero absoluto, cero implica ausencia de contenido.

Tenga en cuenta entregar los ejercicios con el desarrollo completo de su resolución.

A partir de la información presentada en el gráfico:

1. Identifique las unidades de análisis, la población y la o las variables. (0,5 pto.)

Gráfico cuatro:

Establecimientos de educación común que disponen de computadora según tipo de conexión a internet. Total país. En porcentaje.

Fuente: Elaboración propia sobre la base de DINIECE – Relevamiento Anual 2005.

(era un gráfico del tipo "diagrama de sectores" al que le correspondían: telefónica

63,8% - ADSL 15,8% - Cable módem 12,6% - Satelital 5,4% - Sin información 2,4%).

2. Clasifique la o las variables y mencione sus categorías.
3. Reconstruya la distribución en frecuencias absolutas sabiendo que la base es de 23.052 unidades.
4. ¿Cuál es el modo de esta distribución? ¿Cómo clasificaría las entropía de esta distribución?

2.

TABLA 1

X	ni
31 a 35	2
26 a 30	15
21 a 25	37
16 a 20	48
11 a 15	24

En la distribución de frecuencias de la **Tabla 1** se registran los resultados de un grupo de jóvenes (A) en una prueba de historia con 50 ítems.

- a) Calcule la media.
- b) El 45% de los integrantes del grupo que mejores resultados obtuvieron, ¿qué cantidad de ítems respondieron correctamente como mínimo?
- c) Si para otro grupo compuesto por adultos mayores (B) se calculó una media de 30,5 y un desvío estándar de 2,2 ítems; ¿cuál de los dos grupos presenta mayor variabilidad relativa en los resultados? Justifique.
- d) Efectúe un comentario sobre un aspecto evidente de la forma de la distribución.

Teóricas

- a) ¿Qué es la leptocurtosis?
- b) ¿Cuántos tipos de gráficos de barras hay?
- c) Defina al desvío estándar.
- d) ¿En qué consiste a media total o ponderada?

1. ¿Qué es la medición. Diferencias entre nivel de medición ordinal e intervalar.
2. En un ejercicio dado, identificar variables, población/ muestra de individuos y observaciones. Niveles de medición.

3. Medidas de tendencia central. Describalas y diga cuando se utiliza o no cada una.
4. En un ejercicio sacar el D30 y un k.
5. Hay otro ejercicio donde hay que comparar dos grupos sacando media, varianza, desvío, CV. Para uno de ellos ya hay información dada sobre la varianza, desvío típico.

II.

1) En un servicio de psicopatología, de un centro público de salud mental de una Cdad de Salta, se conoce que, la quinta parte de los pacientes abandona el tratamiento antes de culminar el número de sesiones que le indica el psicólogo en la entrevista de admisión de dicha Institución. El próximo mes iniciarán su tratamiento 15 pacientes elegidos al azar.

- a) Mencione la variable Bernoulli presente en este ejercicio y defina la variable binomial. Calcule la probabilidad de:
- b) que abandonen el tratamiento entre 5 y 10 pacientes, antes de cumplir con todas las sesiones?
- c) que a lo sumo 8 no abandonen el tratamiento, antes de cumplir con todas sus sesiones?
- d) que abandone sus sesiones la quinta parte de los pacientes?

2) Estudios anteriores permiten considerar que el puntaje en Mejoramiento Personal de un grupo de estudiantes de bellas artes de las escuelas afines de la Ciudad de Buenos Aires se DN con media 25 y desvío 3,5. Silvio, alumno de primer año obtuyo 27 puntos en M. Personal.

- a) ¿El rango percentilar que le corresponde a Silvio según su puntaje en M. Personal podría ser 30? Responda y justifique, sin hacer cálculos. Si su respuesta es negativa obtenga el rango percentilar que corresponda.
- b) Según el M Personal, otros dos alumnos de primer año, Perla y Carlos, obtuvieron, respectivamente, puntaje $z = 1,6$ y puntaje $T = 40$. Ordene a los tres alumnos por su rendimiento en M. Personal. Justifique su respuesta.

3) Se han tomado dos muestras aleatorias de jóvenes de entre 18 y 25 años de la ciudad de Salta; una, de 30 mujeres y, la otra, de 25 varones. A cada uno de los jóvenes de ambas muestras se le administró la Escala sobre el sentido del humor. Si se sabe que la muestra de mujeres tuvo una media de 39 y un desvío de 4,60, y la de los varones con una media de 35,08, desvío= 4,17

- a) Desarrolle en términos coloquiales las hipótesis planteadas e indique qué tipo de prueba de hipótesis es la exhibida y por qué se utilizó en este caso.
- b) cuales son los supuestos necesarios para llevar a cabo la prueba de hipótesis correspondiente?
- c) con un nivel de significación del 5% que conclusión se desprende de la prueba de hipótesis planteada

PREGUNTAS TEÓRICAS

- 4- Caracterice el modelo de distribución normal. Ejemplifique
- 5- Explique el contraste de hipótesis para una proporción

1.- El rendimiento promedio en un test de habilidad numérica obtenido por los niños de escuelas primarias del Partido de la Matanza de la Prov. De Buenos Aires es de 8,2. Se tomó una muestra aleatoria simple de 41 niños de dicho partido y se les aplicó durante un período de seis meses un nuevo método de enseñanza propuesto por un equipo docente especializado. Al finalizar el período se evaluó al grupo de niños obteniéndose una nota promedio de 9,1 y un desvío de 2,4. ¿Con un nivel de significación del 5%, se puede decir que las diferencias son significativas?

2.- Se aplicó una prueba de memoria auditiva a una muestra de 120 sujetos y arrojó una media de 44 y un desvío típico de 3,5. Suponiendo que los puntajes se distribuyan normalmente:

- a) ¿Cuántos sujetos cabe esperar que obtengan entre 40 y 50 puntos?
- b) ¿Qué puntaje obtuvo el sujeto que fue superado por el 31% del grupo?
- c) ¿Qué porcentaje supera a un sujeto que obtiene 48 puntos?

3.- Sabiendo que el 40% de los niños de la Prov. De Formosa provienen de hogares de con carencias económicas, calcule la probabilidad de que en una muestra aleatoria simple de 8 sujetos:

- a) Por lo menos 1 provenga de hogares con carencias económicas.
- b) Que exactamente 6 provengan de hogares con carencias económicas.
- c) A lo sumo 4 no provengan de hogares con carencias económicas.

4.- Explique el modelo de probabilidad normal. De sus características y propiedades.

Ejercicio 1

El psicólogo laboral de la organización X de la Provincia de Buenos Aires está interesado en estudiar la posible incidencia de los hábitos de ocio extra laboral en el desempeño laboral de los trabajadores de dicha organización en el segundo semestre del año 2013. Para ello, cuando los trabajadores finalizaron sus evaluaciones de desempeño semestral, se les pidió que respondieran a un cuestionario sobre dichos hábitos. El desempeño laboral se evaluó mediante la metodología 360° (sistema de evaluación que tiene en cuenta no solo la evaluación de superiores sino también de compañeros y subordinados) y se obtuvo como resultado un puntaje de 0 a 100. El siguiente cuadro muestra los resultados de la evaluación de Desempeño Laboral de 10 trabajadores y el tiempo semanal que cada cual emplea para el ocio extra laboral.

Nombre	Tiempo de ocio semanal extra laboral	Desempeño Laboral
Sebastián	Menos de 6 horas	40
Valentina	Menos de 6 horas	20
Gabriela	Menos de 6 horas	100
Betiana	Entre 6 y 8 horas	60
Sofía	Entre 6 y 8 horas	50
Guadalupe	Entre 6 y 8 horas	60
María	Entre 6 y 8 horas	40
Juan	Más de 8 horas	80
Esteban	Más de 8 horas	90
Hernán	Más de 8 horas	90

- Detemine la población de individuos y la/s de observaciones.
- Clasifique la variable Desempeño Laboral e indique su nivel de medición.
- Si bien se presentan pocas observaciones para realizar un análisis adecuado ¿Cuál es la fuente sistemática que parece provocar desempeños laborales diferentes?
¿Cuál es la variación que se prevé tendrá la variable bajo estudio por acción de la fuente sistemática identificada?
- ¿Quién no responde a la tendencia prevista?
- A su entender, ¿qué fuente fortuita de variación puede haber incidido para que difieran los desempeños laborales entre sí? Cite una o más.

EJERCICIO 1

En un centro de salud mental especializado en Trastorno de Ansiedad ubicado en el Partido de Quilmes se realizaron estudios que indican que un 12,5% de las consultas son por Fobia específica.

Se sabe que la próxima semana asistirán 15 personas a dicho centro.

Considere la variable "cantidad de pacientes que consultarán por fobia específica entre los 15 que asistirán en la semana próxima" y que el modelo binomial la explica convenientemente

- Indique cuáles son los parámetros de la distribución binomial, su media y desvío.
- A partir de la información que aparece en el siguiente cuadro responda

X	Probabilidad	Probabilidad acumulada	X	Probabilidad	Probabilidad acumulada
0	0,1349	0,1349	8	0,0002	1,0000
1	0,2891	0,4241	9	0,0000	1,0000
2	0,2891	0,7132	10	0,0000	1,0000
3	0,1790	0,8922	11	0,0000	1,0000
4	0,0767	0,9689	12	0,0000	1,0000
5	0,0241	0,9930	13	0,0000	1,0000
6	0,0057	0,9988	14	0,0000	1,0000
7	0,0011	0,9998	15	0,0000	1,0000

- ¿Cuál es la probabilidad de que a lo sumo consulten dos personas por fobia específica?
- ¿Cuál es la probabilidad de que nadie consulte por fobia específica?
- ¿Cuál es la probabilidad de que alguno consulte por fobia específica?
- ¿cuál es la probabilidad de cómo mínimo tres consulten por fobia específica?

EJERCICIO 2

Se sabe que la variable coeficiente intelectual (CI) se distribuye normalmente con media igual a 100 y desvío estándar igual a 15

- Si Ramiro tiene un CI igual a 112 ¿cuál sería su rango percentilar?
- Según sus puntajes en CI, Dante obtuvo un $Z=-1$ y Ximena obtuvo un $T=65$. Ordene a los tres en función de su rendimiento.

X	Media	Desvío	Acumulado	Probabilidad
112	100	15	0	0,019
112	15	100	1	0,834
112	15	100	0	0,002
112	100	15	1	0,788

EJERCICIO 3

Se ha tomado una muestra aleatoria tamaño 43 de los consultantes adultos en un servicio de salud mental en CABA.

Seguidamente, se dividió al grupo entre empleados y desempleados y se les administró el BDI II (inventario de depresión de Beck). Los datos obtenidos fueron procesados con el programa Statistix obteniendo la siguiente salida:

EJERCICIO 1:

A un grupo de estudiantes se les ha administrado una prueba que mide los tiempos de reacción ante determinado estímulo. Considere la variable X definida como "Tiempo de reacción ante el estímulo en la prueba" y suponga que sigue una distribución normal con media de 80 segundos y desvío estándar de 10 segundos. Los resultados de cuatro estudiantes son los siguientes: José Luis obtuvo un puntaje $Z = -0.35$, Pedro obtuvo un puntaje $T = 60$, Juan obtuvo un rango percentilar de 69 y Marcelo obtuvo un tiempo $X = 85$ segundos.

- ¿Quién obtuvo el mejor rendimiento?
- ¿A cuántas desviaciones típicas se encuentra Pedro?
- ¿Cuál es el rango percentilar de Marcelo?

X	MEDIA	DESVIÓ	ACUMULADO	NORMAL
65	80	10	1	0,0899
70	80	10	1	0,1587
75	80	10	1	0,2420
85	80	10	1	0,6915
90	80	10	1	0,8413
95	80	10	1	0,9332
100	80	10	1	0,9772
105	80	10	1	0,9938

EJERCICIO 2:

Los puntajes que obtienen los individuos en una prueba de comprensión lectora se distribuyen normalmente con media 45 y desvío 6. Se considera grupo de riesgo al formado por los alumnos que obtienen menos de 43 puntos en la prueba. Si se administra la prueba a 1000 individuos:

- ¿Cuántos de ellos se espera que queden en situación de riesgo?
- ¿Cuál es el puntaje que deja por encima aproximadamente al 75 % de la población?
- ¿Cuál es la probabilidad de que un alumno supere un puntaje de 42 puntos pero obtenga a lo sumo 51 puntos?

x	media	desvío	acumulada	normal
---	-------	--------	-----------	--------

1. Se desarrolló un programa de prevención de la violencia en una escuela de esta ciudad. Se asignaron 40 niños de manera aleatoria a dos grupos: los niños del grupo 1 participaron en talleres sobre dicha temática y sus padres recibieron charlas orientadoras; los niños del grupo 2 participaron de los mismos talleres pero sus padres no fueron incluidos en las charlas. Luego de 2 meses, se midieron los niveles de violencia de los niños mediante una escala y se realizó una prueba de hipótesis. Se presentan los datos obtenidos con el Statistix.

Two-Sample T Tests for G1 vs G2

Variable	Mean	N	SD	SE
G1	12,750	20	2,5314	0,5660
G2	14,950	20	2,4597	0,5500
Difference	-2,2000			

Null Hypothesis: difference = 0

Alternative Hyp: difference < 0

Assumption	T	DF	P	95% CI for Difference	
				Lower	Upper
Equal Variances	-2,79	38	0,0041	-3,7977	-0,6023
Unequal Variances	-2,78	38.0	0,0039	-3,7978	-0,6022

Test for Equality of variances	F	DF	P
	1,06	19,19	0,4508

- A) *Ponga en palabras el planteo expresado en las hipótesis, considerando la variable en juego*
- B) *Marque en la hoja el valor del estadístico de prueba utilizado. Justifique*
- C) *Suponiendo la normalidad de la distribución de los puntajes en la escala de violencia, lleve adelante la prueba de hipótesis correspondiente, utilizando un nivel de significación del 5%*

2. Se desea investigar si existe relación entre el número de sucesos vitales estresantes sufridos en el último año por un sujeto y el puntaje en un test de rendimiento. A continuación se dan los resultados obtenidos:

Unweighted Least Squares Linear Regression of Puntaje en el test de rendimiento					
Predictor					
Variables	Coefficient	Std Error	T	P	
Constant	8,31092	0,70872	11.73	0	
Nº sucesos estrés	-1,81513	0,41453	-4,38	0,0011	
R-Squared	0,6354	Resid. Mean Square (MSE)		3,14591	
Adjusted R-Squared	0.6023	Standard Deviation		1,77367	
Source	DF	SS	MS	F	P
Regression	1	60,3180	60,3180	19,17	0,0011
Residual	11	34,6050	3,1459		
Total	12	94,9231			

Correlations (Pearson)	
Puntaje en el test de rendimiento	Nº de sucesos estresantes
	-0,7971
Cases	
Included 13	Missing Cases 0

- A) *Determinar mediante un coeficiente adecuado el grado de ajuste de los datos observados al modelo lineal. Justifique.*
- B) *Determinar mediante un coeficiente adecuado el grado y sentido de la asociación entre ambas variables. Interpretar su resultado. Justifique.*
- C) *Hallar la recta de regresión para predecir el puntaje en el test de rendimiento a partir de la cantidad de sucesos estresantes vividos.*
- D) *Predecir el puntaje en el test de rendimiento para un sujeto que vivió 4 sucesos estresantes en el último año.*

3. Se desea estudiar la relacion entre "estado civil de un sujeto" (X) y "estado civil de la madre" (Y). Utilizando los datos que se presentan a continuación, realice la prueba correspondiente, con un nivel de significación del 1%.

Chi-Square Test for Heterogeneity or Independence					
for 1 = H X					
Y		X			
		Soltero	Casado	Separado	
Soltero	Obs	3	1	2	6
	Expected	2,75	1,25	2,00	
	Cell Chi-Sq	0,02	0,05	0,00	
Casado	Obs	6	1	3	10
	Expected	4,58	2,08	3,33	
	Cell Chi-Sq	0,44	0,56	0,03	
Separado	Obs	2	3	3	8
	Expected	3,67	1,67	2,67	
	Cell Chi-Sq	0,76	1,07	0,04	
		11	5	8	24
Overall Chi-Square		2,97			
P-Value		0,5623			
Degrees of Freedom		4			

Preguntas Teóricas:

- 1) Explique el carácter variable de la media muestral ¿Qué significa que la misma es un estimador insesgado de la media poblacional?
- 2) ¿A que se hace referencia con la expresión "marco de muestreo"? Explique que requisitos debe cumplir este marco.
- 3) ¿Cómo se refleja una correlación lineal débil o nula entre dos variables en el diagrama de dispersión?
- 4) Explique la noción de relación entre sucesos ¿Por qué interviene la estadística en el estudio de las relaciones entre observaciones?