

Tema 1

Estadística Descriptiva

Estadística: Es la parte de las matemáticas que trata de obtener Estadística a partir de unos datos numéricos.

- Estadística descriptiva: se encarga de la recogida, organización y presentación de los datos obteniendo una serie de parámetros que los representen. Es totalmente objetiva.
- Estadística inferencial: se encarga del análisis y toma de decisiones a partir de los datos anteriores. Es un poco subjetiva. Hace uso del cálculo de probabilidades.

Definiciones:

- Llamamos población a un conjunto de elementos con una característica común.

A cada elemento de la población, lo llamamos individuo.

- Llamamos muestra a cualquier conjunto de la población.

Al número de individuos que componen la muestra se los llama tamaño muestral.

- Llamamos variable estadística a cualquier característica de los individuos de la población que podemos analizar.
- Variables categóricas o cualitativas: son aquellas que no se pueden cuantificar numéricamente.
- Variables cuantitativas: aquellas que se pueden cuantificar numéricamente. Se pueden clasificar en:
 - ♦ Discretas: aquellas que pueden tomar un número finito o infinito numerable de valores.
 - ♦ Continuas: aquellas que pueden tomar todos los infinitos valores de un intervalo.

Definición: Llamamos frecuencia absoluta n_i de un valor X_i al número de veces que aparece dicho valor.

tamaño muestral Suponiendo $X = \{x_1, x_2, \dots, x_m\}$

Llamamos frecuencia absoluta acumulada N_i de un valor X_i al número de veces que aparece dicho valor y todos los anteriores a él.

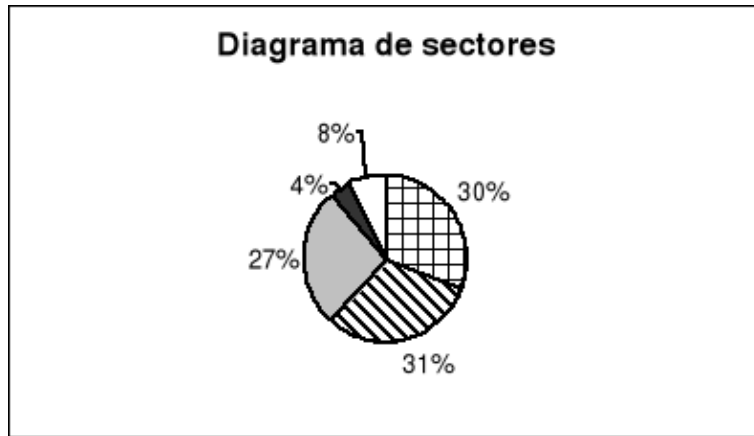
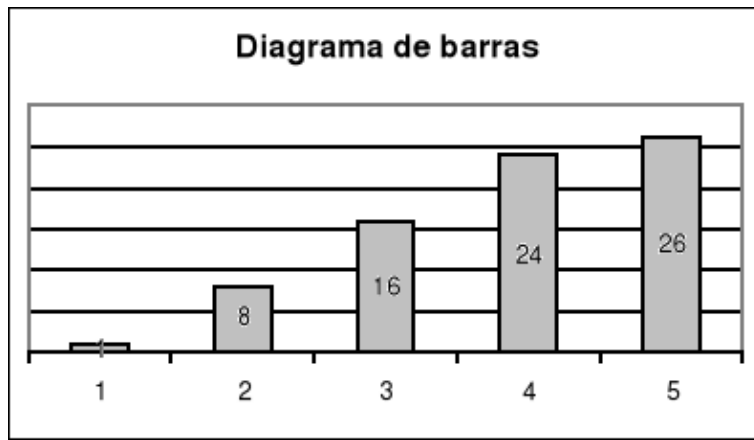
Definición: Llamamos frecuencia relativa f_i de un valor X_i al cociente de la frecuencia absoluta de X_i y el tamaño de la muestra.

Llamamos frecuencia relativa acumulada F_i de un valor X_i a la suma de las frecuencias relativas de dicho valor y de los anteriores a él.

Observación: Si $f_i \cdot 100$ obtenemos el % de incidencia del valor X_i sobre la muestra.

Representaciones gráficas:

Ejemplos: (OA)



Para variables continuas o discretas con muchos valores, los datos se agrupan en clases.

Para formar las clases:

- Redondeamos los valores (depende de los datos y precisión necesaria)
- Elegimos el número de clases (entre 5 y 20)
- Escribimos las clases de manera que cada elemento esté perfectamente situado en 1 y solo 1 de ellas.

Al punto medio se le llama marca de clase X_i

marca de clase

Diagrama de tallo y hojas:

Ejemplo: Altura

Histograma:

Lo que se debe ver en un histograma es:

- Casos atípicos: son casos que son raros en función de los datos.
- Simetría: se observa si el histograma es uniforme o simétrico en alguna de sus partes.

Medidas asociadas a una variable:

- Centralización:

- Media aritmética:

Características:

- Tiene en cuenta todos los datos: es muy representativa.
- Es muy sensible a los datos atípicos y asimétricos.
- No se puede calcular para variables categóricas.
- Moda (Mo): Es el valor de la variable que más se repite.
 - Es la única medida de centralización que se puede obtener para variables categóricas.
 - Cuando tengamos clases hablaremos de clase modal.
 - Si dos valores tienen frecuencia máxima hablaremos de variable bimodal.
 - Mediana (Me): Es un valor de la variable que deja el 50% de la distribución ordenada a su derecha y el otro 50% a su izquierda.
 - No tiene en cuenta los valores de la variable, sino su situación.
 - No es nada sensible a datos atípicos.
 - Cuando trabajamos con clases, hablamos de clase mediana.
 - Cálculo:
Si el n° de datos es impar, tomamos el elemento x_i con

Si el n° de datos es par, la Me es la semisuma de los dos datos centrales.

- Medidas de dispersión:

- Rango o recorrido:
- Varianza:

- Tiene en cuenta todos los valores.
- Es muy sensible a los valores atípicos y asimétricos.
- Se expresa en las unidades de la variable al cuadrado.
- Desviación Típica:

Observación:
Se estima que en el

intervalo se encuentra el 68% aproximadamente de los valores de la variable.

En el intervalo 95%

En el intervalo 99,7%

· Rango intercuartílico:

Cuartiles:
Son 3 valores $Q1$, $Q2$ y $Q3$ que dividen a la distribución ordenada en 4 partes iguales.

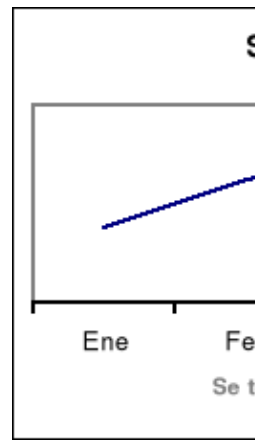
$Q2 = Me$;

$Q1 = Me$ de los valores situados a la derecha de $Q2$;

$Q3 = Me$ de los valores situados a la izquierda de $Q2$.

Definición:
Coeficiente de variación:

Series temporales:



Distribuciones

Bivariantes:

Suponemos
que x e y
son dos
variables
sobre la
misma
población.

- x

toma

valores

$x_1, x_2, \dots, x_n$

(x, y) toma

valores

$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

- y

toma

valores

$y_1, y_2, \dots, y_n$

Covarianza:

Siendo n_i la
frecuencia
de (x_i, y_i)

La
covarianza
relaciona
los
parámetros
 x e y .

Coefficiente de correlación:

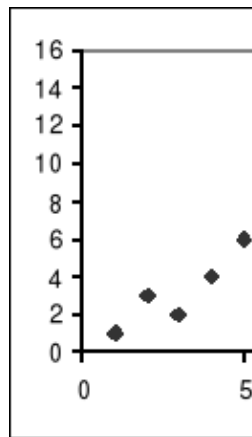
adimensional

$-1 \leq r \leq 1$

Diremos
que hay una
mayor
relación
entre las 2
variables
cuando r se
vaya
acercando a
los
extremos,
mientras
que si está
cerca del 0
se dice que
hay poca
relación
entre las 2
variables.

Ajuste por
mínimos
cuadrados:

Recta de
regresión de
y sobre x



Funciones
linealizables:

Función
exponencial:

Modelo
potencial:

Tema 2

Fundamentos
de la Teoría
de la
Probabilidad

Definición:
Llamamos
experimento
aleatorio o
estocástico
a aquel que
realizado en
idénticas
circunstancias
nunca
podemos
predecir los
resultados.
Como por
ejemplo:
lanzar un
dado;
extraer una
carta.

En caso
contrario el
experimento
se llama
determinista.
Por
ejemplo:
caída libre.

Definición:
A cada una
de las veces
que
repetimos
un
experimento
aleatorio la
llamamos
prueba.

Definición:
Al conjunto

de los
posibles
resultados
de un
experimento
aleatorio se
le llama
espacio
muestral
(E).

Definición:
Llamamos
suceso a
cualquier
subconjunto
del espacio
muestral.
En el
ejemplo de
lanzar un
dado:

$$E=\{1,2,3,4,5,6\}$$

Sucesos:
 $A=\{\text{salir } 1\}$

$B=\{\text{salir}$
 $\text{par}\}=\{2,4,6\}$

$$C=\{1,2,3,4,5,6\}=E$$

$$D=\{\}$$

Tipos de
sucesos:

- Suceso elemental:
elemental:
compuesto
por
un
solo
elemento
muestral.
Ejemplo:
 A
- Suceso compuesto:
compuesto:
formado
por

2 o
más
puntos
muestrales.

Ejemplo:

B

- Suceso
cierto
o
seguro:
siempre
se
verifica
coincide
con
el
espacio
muestral.

Ejemplo:

C

- Suceso
imposible:
nunca
se
verifica
".

Ejemplo:

D

Definición:

Dado un
suceso
cualquiera
A,
llamamos
suceso
opuesto,
contrario o
complementario
ó a aquel
que se
verifica
cuando no
se verifica
A.

Ejemplo:
lanzar un
dado

$A = \{\text{salir par}\} = \{2, 4, 6\}$

= {salir
impar} = {1,3,5}

Operaciones
con
sucesos:

- Inclusión:
diremos
que
 A
está
incluido
en
 B , o
que
 A
implica
 B si
cuando
se
verifica
 A
entonces
también
se
verifica
 B .
Ejemplo:

$B = \{\text{salir}$
 $\text{par}\};$
 $A = \{\text{salir } 4\}$

- Igualdad:
diremos
que
 A es
igual
a B
si A
está
incluida
en
 B y
 B
está
incluida
en
 A
($A=B$).
Ejemplo:

$A = \{\text{salir}$

par}

$B = \{\text{salir}$
 $2, 4, 6\}$

- Unión:
dados
2
sucesos
 A y
 B ,
llamamos
unión
de
 A y
 B a
otro
suceso
que
se
verifica
si
se
verifica
 A o
se
verifica
 B .
Ejemplo:

$A = \{\text{salir}$
 $\text{par}\} = \{2, 4, 6\}$

$B = \{\text{salir}$
 $\text{primo}\} = \{2, 3, 5\}$

$= \{2, 3, 4, 5, 6\}$

- Intersección:
dados
2
sucesos
 A y
 B ,
llamamos
intersección
de
 A y
 B a
otro
suceso
que
se

verifica
si
se
verifica
 A y
se
verifica
 B .

Ejemplo:

$A = \{\text{salir par}\} = \{2, 4, 6\}$

$B = \{\text{salir primo}\} = \{2, 3, 5\}$

$= \{2\}$

Definición:

Diremos
que A y B
son sucesos
incompatibles
si $=$ " (no se
verifican
simultáneamente)

- Diferencia:
dados
2
sucesos
 A y
 B ,
llamamos
diferencia
de
 A y
 B a
otro
suceso
que
se
verifica
cuando
se
verifica
 A y
no
se
verifica
 B .

Ejemplo:

$A = \{\text{salir par}\} = \{2, 4, 6\}$

$B = \{\text{salir primo}\} = \{2, 3, 5\}$

$A - B = \{4, 6\}$

$B - A = \{3, 5\}$

- Diferencia simétrica:
dados
2
sucesos
 A y
 B ,
llamamos
diferencia
simétrica
de
 A y
 B a
otro
suceso
que
se
verifica
 A y
no
se
verifica
 B ó
se
verifica
 B y
no
se
verifica
 A .

Ejemplo:

$A = \{\text{salir par}\} = \{2, 4, 6\}$

$B = \{\text{salir primo}\} = \{2, 3, 5\}$

Definición:
Diremos
que n
sucesos de

un espacio
muestral E ;
 $A_1, A_2, \dots, A_n$
forman un
sistema
completo de
sucesos si:

- - Propiedades
- Conmutativas:
- Asociativas:
- Idempotentes:
- Distributivas:
- Leyes de
Morgan:

Introducción al concepto de probabilidad

Probabilidad
es la medida
de la
incertidumbre
asociada a
un suceso

Definición:
Llamamos
probabilidad
de un
suceso A al
valor
entorno al
cual tiende
a
estabilizarse
la
frecuencia
relativa del
suceso A
cuando se
repite un n°
suficiente
de veces.

Definición axiomática de Kolmogorov:

Definición:
 Llamamos función de probabilidad a una aplicación del espacio de sucesos $(\Omega, \mathcal{F})$.

que cumple
 3 axiomas:

Consecuencias:

·
 En efecto:

·
 En efecto:

·
 En efecto:

·
 En efecto:

· Si E está compuesto por n sucesos elementales $A_1, A_2, \dots, A_n$ equiprobables entonces

En efecto:

· Regla de Laplace: Si E está compuesto por n sucesos elementales equiprobables y sea A un suceso compuesto por k sucesos, entonces:

conteo:

- Combinatori

disponibles

elegidos

¿Influye el orden de elección?

condicionada

Definición:

a B a:

Definición:

que A y B

son
independientes
si
 $P(A)=P(A/B)$

En caso
contrario se
dicen
dependientes.

Proposición:
A y B son
independientes
si y solo si

En efecto:

Definición:
Experimentos
compuestos:
aquel que
esta
formado por
2 o más
experimentos
compuestos.

Teorema de
la
probabilidad
compuesta:

Sean
 $A_1, A_2, \dots, A_n$;
n sucesos
tales que

Entonces:

Teorema de
la
probabilidad
total

Sean
 $A_1, A_2, \dots, A_n$;
un sistema
completo de
sucesos con

Sea B un
suceso

cualquiera.

Entonces:

Demostración:

Teorema de Bayes

En las mismas condiciones del teorema anterior

Tema 3

Variables Aleatorias I

Definición:

Una variable aleatoria es una función que asocia a cada elemento del espacio muestral un número real.

Observación:

Sobre un mismo espacio muestral se pueden definir varias variables aleatorias.

Ejemplo:

Experimento:
Lanzar un dado

$E = \{1, 2, 3, 4, 5, 6\}$

Observación:

Si el resultado del experimento son valores numéricos se les suele asociar el mínimo valor.

Si son resultados cualitativos se les asocia cualquier número.

Ejemplo:

Experimento:
Lanzar moneda

$E = \{C(\text{cara}), +(cruz)\}$

Exp: Lanzar
3 monedas

$E = \{CCC, CC+, C+C, C++\}$

Definición:

Función de distribución

es una función que asocia a cada valor de la variable aleatoria la probabilidad acumulada hasta dicho valor.

Ejemplo:

Propiedades:

- $F(t)$

es
creciente
• $F(t)$
es
continua
por
la
derecha

**Variables
Aleatorias
Discretas:
Función
Puntual de
Probabilidad**

Definición:

Diremos
que una
variable
aleatoria X
es discreta
si puede
tomar un
numero
finito de
valores o
infinito
numerable.

Definición:

Llamamos
función
puntual de
probabilidad
a una
función que
asocia a
cada valor
de la
variable su
probabilidad.

Propiedades:

Nota: estas
2
propiedades
caracterizan
una función
puntual de
probabilidad.

· Ordenando los valores

Características de una Variable Aleatoria Discreta

Definición:
Sea X una v.a. discreta con función puntual de probabilidad llamamos esperanza matemática de la v.a. X a:

Observación:
La existencia de $E(X)$ depende de la convergencia de la serie cuando X toma un n° infinito numerable de valores.

Propiedades de la esperanza:

Definición:
Sea X una v.a discreta con función puntual de probabilidad . Llamamos varianza de la v.a. X a:

siendo

Propiedades:

Demostración:

Propiedades:

Definición:

Llamamos desviación típica de la v.a. X a la raíz cuadrada positiva de la varianza

Variables Aleatorias Continuas: Función de Densidad

Definición:

Una v.a es continua cuando puede tomar todos los valores de un intervalo.

Ejemplo:

Altura de las personas, longitud de un tornillo, etc.

Observación:

No se puede calcular la probabilidad de una valor determinado ya que siempre vale 0. Entonces calculamos la probabilidad de que la

variable se
sitúe en un
intervalo.

Definición:
Supongamos
que
tomemos
una v.a.
continua y
los valores
que toma
los
distribuimos
en clases.
Entonces
podemos
hacer el
histograma
de esa
variable.

Llamamos
función de
densidad
 $f(X)$ a la
función
hacia cuya
gráfica
tienden los
histogramas
de la v.a.
cuando
elegimos las
clases cada
vez mas
pequeñas.

Propiedades:

Características de una Variable Aleatoria Continua

Definición:
Sea X una
v.a.
continua
con función
de densidad

$f(X)$.

Llamaremos
esperanza
de X ó a:

Observación:
La
existencia
de $E(X)$
depende de
la
convergencia
de la
integral.

Definición:
Sea X una
v.a.
continua
con función
de densidad
 $f(X)$.
Llamamos
varianza de
 X ó a:

Definición:
Llamamos
desviación
típica de X
a:

Distribuciones Asociadas a Variables Aleatorias Discretas

· Distribución
de
Bernoulli:
Suponemos
que un
experimento
aleatorio
tiene 2
posibles
soluciones

Función
puntual de
probabilidad
de la
distribución
de
Bernouilli

· Distribución
Binomial:
Suponemos
que
repetimos
un
experimento
un numero
n de veces
de manera
independiente.
Si
definimos:

$X =$ nº de
éxitos
obtenidos
en los n
experimentos

Diremos
que X sigue
una
distribución
Binomial de
parámetros.

Función
puntual de
probabilidad
de la
distribución
Binomial

Es función
puntual de
probabilidad:

X la
podemos
considerar
como suma
de n
variables X_i

que siguen
una
distribución
de
Bernouilli.

· Distribución
de Poisson:
Suponemos
que
estudiamos
un suceso
que ocurre
esporádicamente
a lo largo de
un soporte
continuo

$X = n^\circ$ de
veces que
aparece
dicho
suceso en
un periodo
de tiempo

Diremos
que X sigue
una
distribución
de Poisson
de
parámetro

Si se
cumplen 2
condiciones:

- El suceso
aparece, a
largo plazo,
un n° medio
de veces
- La
aparición
del suceso
ocurre de
manera
independiente
(el n° de
veces que
ocurre en un

intervalo de tiempo no induce a saber el n° de veces que ocurrirá en el siguiente intervalo).

Es función puntual de probabilidad:

Observación:
Tanto la distribución Binomial como la Poisson son reproductivas respecto a los parámetros n y p respectivamente.

Distribuciones Asociadas a Variables Aleatorias Continuas

Distribución Normal:
Una v.a. X sigue una distribución normal de parámetros

μ y σ^2 tiene por función de densidad

Observación:

Tipificación de la normal

Si Z defino

Ejemplo:

Teorema de Moivre-Laplace

Si entonces

Explicación:

Si

Observación:

La distribución de Poisson se puede aproximar a una normal cuando $n > 5$

Si

Distribución exponencial:

Son variables asociadas a tiempos de espera para recibir determinados servicios. X sigue una distribución exponencial de parámetros

Si tiene por función de densidad

Función de distribución

Esperanza

Varianza

Observación:

Tanto la

distribución
normal
como la
exponencial
son
reproducidas
respecto a
sus
parámetros.

Tema 4

Variables Aleatorias II

Definición:
Si X e Y son
dos v.a.
definidas
sobre el
mismo
espacio
muestral.
Llamamos
variable
aleatoria
bidimensional
 (X,Y) a una
función

Definición:
Llamamos
función de
distribución
de una
variable
aleatoria
bidimensional
 (x,y) a

Variables Aleatorias Bidimensionales Discretas

Si llamo C
al conjunto
de posibles
resultados
de la v.a.b.
 (X,Y) .

Si C tiene una cantidad finita o infinita numerable (x,y) se llama v.a.b. discreta.

Definición:
Llamamos función puntual de probabilidad conjunta de una v.a.b. discreta a una función:

con

.
Variables
Aleatorias
Bidimensionales
Continuas

Si C tiene una cantidad infinita de puntos (X,Y) se llama v.a.b. continua.

Definición:
Llamamos función de densidad conjunta de una v.a.b. continua:

con

.
Distribuciones
Marginales

Si (x,y) es una v.a.b. con función de distribución $F(x,y)$. Llamamos funciones marginales de distribución de x e y a:

- si la v.a. es discreta

función puntual de probabilidad de X

función puntual de probabilidad de Y

- si la v.a. es continua

función de densidad marginal de X

función de densidad marginal de Y

Independencia de Variables

Definición:
Si (x,y) es una v.a.b. diremos que X e Y son independientes si la función

de
distribución
conjunta es
el producto
de las
marginales

- si la
v.a.
es
discreta

- si la
v.a.
es
continua

Observación:
Si X e Y son
independientes

Tema 5

Muestreo y
Distribuciones
Muestrales

Definición:

Población:
es un
conjunto de
elementos
con una
característica
común (que
será el
objeto de
estudio).

Individuo:
es cada uno
de los
elementos
de la
población.

Muestra: es
cualquier
subconjunto
de la
población.

Tamaño muestral: es el numero de elementos que componen la muestra.

La característica principal de una buena muestra es que sea representativa de la población a estudiar.

Definición:
Muestreo Aleatorio Simple (M.A.S.) es el que obtiene la muestra teniendo en cuenta:

- Todos los individuos de la población tienen la misma posibilidad de ser elegidos.
- Las extracciones han de realizarse con reemplazamiento.

Si la población es muy grande el reemplazamiento

no tiene
importancia.

Otros tipos
de
muestreo:

- Muestreo
sistemático:
seleccionamos
el
primer
elemento
de
manera
aleatoria
y
los
demás
sistemáticamente
como
se
halla
fijado.
- Muestreo
estratificado
- Muestreo
por
conglomerados

Estadísticos
muestrales

Suponemos
que
tenemos
una v.a. X
definida
sobre una
determinada
población.
Sea
 $(X_1, X_2, \dots, X_n)$
una m.a.s.
de tamaño
 n . A cada
 $(X_1, X_2, \dots, X_n)$
se le llama
realización
de la
muestra.

Si llamamos

El conjunto de todas las posibles realizaciones de la muestra; llamamos estadístico muestral a toda función real:

Observación: como un estadístico es función de variables aleatorias este también será variable aleatoria con función de densidad

por ser X_i independientes e idénticamente distribuidas.

Distribución de la media muestral:
Teorema central de límite

Suponemos que X una v.a. que sigue una distribución normal de media μ , y varianza σ^2 . Tomamos $(X_1, X_2, \dots, X_n)$ una m.a.s. y

definimos

Varianza:

Si

$X \sim N(\mu, \sigma^2)$

Teorema
central de
límite

Si X es una
v.a.

cualquiera
con media
y varianza

y
 $(X_1, X_2, \dots, X_n)$

es una
m.a.s. de
tamaño n

Observación:

n es
suficientemente
grande si
 $n \geq 30$

Observación:

si no
conocemos
la varianza
poblacional
de la v.a.
 X entonces
sigue una
distribución
que se llama
T de Student
con $n-1$
grados de
libertad.

Distribución
asociada a
la
proporción
muestral

Suponemos
que p es la
proporción

de
individuos
de una
población
de una
determinada
característica.
Si tomamos
una muestra
aleatoria
simple de
tamaño n y
denotamos
a la
proporción
muestral

Graficas de control

Definición:
Diremos
que un
proceso esta
bajo
control. Si
la variable
que lo
describe
sigue la
misma
distribución
a lo largo
del tiempo.

Definición:
Las gráficas
de control
son
instrumentos
que se
utilizan para
ver si un
proceso se
encuentra
bajo
control.

Pasos a
seguir para
hacer una
gráfica de
control

- Definimos la variable que nos define el proceso
- Tomamos muestras de pequeño tamaño a lo largo del tiempo
- Dibujar una línea central en la media

Dibujar las
2 líneas de
control

- Representamos en la gráfica los puntos correspondientes

a las
medidas de
todas las
muestras
· Ver si el
proceso está
bajo
control.

Estará bajo
control si:

- Algún
punto queda
fuera de los
límites de
control, el
proceso está
fuera de
control
- Si aparecen
9 o más
puntos
seguidos
entre la
línea media
y una de las
2 líneas de
control
- Si 2 de cada
3 puntos
están
situados
entre las
líneas de
control y las
líneas está
fuera de
control

Tema 6

Introducción
a la Teoría
de la
Aproximación

Estimación
Puntual

Definición:
Un

estimador
es un
estadístico
que intenta
aproximar
el verdadero
valor de un
parámetro
poblacional
desconocido.

- Ejemplo:
donde
 θ es
desconocido

Tomamos
una m.a.s.
de tamaño n

es un
estimador
de θ

X^2 es un
estimador
de m

Propiedades
de los
estimadores:

Definición:
Diremos
que un
estimador
es centrado
ó insesgado
se la
esperanza
del
estadístico
coincide
con el
parámetro
poblacional
estimado.

= parámetro
poblacional
(desconocido)

= estimador

Observación:

se llama
sesgo

Definición:

Llamamos
eficiencia o
precisión de
un
estimador a
la inversa
de la
varianza

Precisión

Observación:

- Si tenemos
2
estimadores
con
propiedades
contradictorias.
Entonces
tomamos el
que menor
error
cuadrático
medio
tenga,
siendo:
- Si aún así
no nos
satisface
exigiremos
que al
menos sea
consistente

Estimación

por
intervalos
de
confianza

Definición:

Llamamos
intervalo de
confianza
para un
parámetro

poblacional
desconocido
con nivel de
confianza
 $1-\delta$ a una
expresión
del tipo de
manera que

Observación:

Esto quiere
decir que en
el
 $100.(1-\delta)\%$
de los
intervalos
que
fabriquemos
con las
muestras
obtenidas
de la
población,
el verdadero
valor del
parámetro
se encuentra
dentro del
intervalo.

Intervalo de
confianza
para la
media
poblacional
con ____

Supongo
que con
conocido y
queremos
fabricar un
intervalo de
confianza
para δ con
nivel de
confianza
 $(1-\delta)$.

Tomamos
una m.a.s.
de tamaño

n ; es un
buen
estimador
puntual de
 θ , y
sabemos
que

Intervalo de
confianza
para θ con

conocido
para X
normal

Determinación del tamaño muestral

Error
máximo
cometido al
estimar θ
mediante
 $\bar{X}$
es:

para una
confianza
dada $1 - \alpha$

Por tanto
fijado el
error:

es el tamaño
muestral
para
cometer
como
máximo
dicho error

Ejemplo: El
tiempo de
reacción de
un motor
sigue una
distribución
 $N(\theta, 0.052)$
¿Cuántas

medidas
deben
hacerse para
que al nivel
de
confianza al
95% error
de la
estimación
 $\bar{X}$
no sea
mayor de
0.01?

$$E = |\bar{X} - \mu| = Z_{1-\frac{\alpha}{2}}$$

Datos:

$$E = 0,01$$

$$\sigma = 0,05$$

$$n = ? \quad 0,1 =$$

$$Z_{1-\frac{\alpha}{2}} = 1,96$$

Por tanto la
muestra
debe ser:
 $n = 97$

Tema 7

Introducción a los Contrastes

Definición:
Una
hipótesis
estadística
es una
proposición
(afirmación)
sobre un
parámetro
(desconocido)
de una o
varias
poblaciones.

Definición:
 Un test o
contraste de
hipótesis es
 un
 procedimiento
 que nos
 hace tomar
 decisiones
 sobre
 determinadas
 proposiciones.

Definición:
Hipótesis
nula H_0 que
 es la que
 suponemos
 cierta de
 partida

Definición:
Hipótesis
alternativa
 H_1 que es la
 que
 suponemos
 cierta
 cuando
 rechazamos
 H_0 .

Definición:
 Error tipo I
 es el que
 cometemos
 al rechazar
 H_0 siendo
 ésta cierta.

Definición:
 Error tipo II
 es el que
 cometemos
 al aceptar
 H_0 siendo
 H_1 cierta
 (H_0 falsa).

<i>Decisión</i>	<i>H₀</i> <i>cierta</i>	<i>H₁</i> <i>c</i>

Acepto H_0	Decisión correcta	E ti
Rechazo H_0	Error tipo I	D c

Observación:

Lo ideal
sería que
 $P(\text{Error tipo I}) = P(\text{Error tipo II}) = 0$

Definición:

Llamamos
nivel de
significación
 δ un n° fijo
a de manera
que $P(\text{Error tipo I}) = \delta$

Definición:

Llamamos
potencia del test a la
probabilidad
de rechazar
 H_0 siendo
 H_1 cierta

Decisión	H_0 cierta	H_1 cierta
Acepto H_0	Decisión correcta $1 - \delta$	E ti $1 - \delta$
Rechazo H_0	Error tipo I δ	D c δ

Observación:

- La potencia nos mide la fiabilidad del

test.
 Consideramos
 fiable
 si la
 potencia
 es
 superior
 a
 0,8
 • El
 nivel
 de
 significación
 δ , si
 no
 lo
 tenemos,
 suponemos
 por
 defecto
 que
 es
 $\delta=0,05$

Procedimiento
 a seguir
 para un test
 de hipótesis

- Seleccionamos
 el
 parámetro
 que
 queremos
 contrastar
- Planteamos
 la hipótesis
 nula $H0$
- Planteamos
 la hipótesis
 alternativa
 $H1$
- Seleccionamos
 el nivel de
 significación
- Elegimos
 un
 estadístico
 adecuado
- Distinguimos
 entre la
 región de
 aceptación

y rechazo
para dicho
estadístico
· Calculamos
el valor de
dicho
estadístico
bajo la
hipótesis
nula
· Tomamos
decisiones
sobre H_0

Contraste
para la
media de
una
población
normal ____
conocida

Suponemos
con
varianza
conocida

· Queremos
hacer un
contraste
para θ
· Tomamos θ
· Sabemos
que es un
buen
estimador
de θ

El
estadístico
adecuado
será

.

.

· Rechazamos H_0 si

No tenemos argumentos para rechazar H_0 si

Definición:
p-valor es el menor nivel de significación para el que rechazamos H_0

Test unilaterales:

Test bilaterales:

Cálculo de la potencia del test

Suponemos m.a.s.

Se mide la fiabilidad del test

Observación:
Para calcular la potencia tenemos que tener una alternativa fija

Suponemos que la alternativa $\delta=6$

$1-\delta = P(\text{Rechazar})$

H_0 / H_1
 $cierta) = "$

Observación:
El test será
fiable si

Segunda
forma de
calcular la
potencia

Calculamos

Z_0 No sigue
una $N(0,1)$

Z_0 – Sigue
una $N(0,1)$

Observación:
En el caso
de contraste
bilateral
como hay 2
zonas de
rechazo,
entonces
calculamos
 $\tilde{\alpha}$

Tema 8

Inferencia
para la
Media

Suponemos
que
queremos
estimar la
media de
una
población
en la que la
varianza sea
desconocida,

· Si n es
suficientemente
grande
aproximamos

(varianza
poblacional)
por medio
de S^2

(varianza
muestral) y
aplicamos
lo anterior:

$$Z = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim N(0, 1)$$

· Si n no es
suficientemente
grande
entonces
usamos el
estadístico:

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

t_{n-1} =
distribución
t de Student
con $n-1$
grado de
libertad

Intervalo de
confianza:

Contraste
de
hipótesis:

$X \sim N(\mu, \sigma^2)$
con
desconocida:
tomamos
una m.a.s.
de tamaño
 $n < 30$

Rechazo H_0
si

Acepto H_0
si

Inferencia
para dos

medias:

Suponemos
que
tenemos dos
poblaciones

con $\hat{E}$
conocida;
tomamos
una m.a.s.
de tamaño
 n_1

con $\hat{O}$
conocida;
tomamos
una m.a.s.
de tamaño
 n_2

$$\bar{X}_1 \sim N \left(\mu_1, \frac{\sigma_1^2}{n_1} \right)$$

$$\bar{X}_2 \sim N \left(\mu_2, \frac{\sigma_2^2}{n_2} \right)$$

Estadístico:

$$\frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

Inferencia
para dos
poblaciones
con
varianzas
poblaciones
desconocidas
(distintas)

Suponemos
que
tenemos dos
poblaciones

$X_1 \sim N(\mu_1, \sigma_1^2)$
con $\hat{E}$
conocida;

tomamos
una m.a.s.
de tamaño
 n_1

$X_2 \sim N(\mu_2, \sigma_2^2)$
con $\hat{O}$
conocida;
tomamos
una m.a.s.
de tamaño
 n_2

Estadístico:
Aproximamos
 $\hat{E}$ por S_{12}
y $\hat{O}$ por
 S_{22}

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$k = \min\{n_1 - 1, n_2 - 1\}$$

Ejemplo:
Altura

$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

Se puede
representar
como un eje
de
coordenadas

$$y = ax + b$$

La ideal
recta que se
busca es la
que hace
mínima la
distancia a
todos los
puntos, pero
al ser difícil
de
encontrar:

1.

2.

3. Es
simétrica
respecto de
 $=Me=Mo$

4. Tiene
puntos de
inflexión en
- y +

5. $f(x)$ no
tiene
primitiva

- +

Grafica de control

Cualquiera
de los 2 es
válido pero
 X se
aproximará
mas al valor
 δ

$\delta =$
 $P(\text{Rechazar}$
 H_0 / H_0
cierta)

$1 - \delta =$
 $P(\text{Aceptar}$
 H_0 / H_0
cierta)

$\delta =$
 $P(\text{Aceptar}$
 H_0 / H_1
cierta)

$1 - \delta =$
 $P(\text{Rechazar}$
 H_0 / H_1
cierta)

Fórmula

General de
 k

$$k = \frac{\frac{S_1^2}{n_1} + \frac{1}{n_1 - 1} \frac{S_1^2}{n_1} + \dots}{\frac{S_1^2}{n_1} + \frac{1}{n_1 - 1} \frac{S_1^2}{n_1} + \dots}$$