FORMULAS IMPORTANTES

ESTADÍSTICA DESCRIPTIVA

VARIANZA: D. TIPICA=

COEF. VARIACIÓN =

DISTRIBUCIONES BIDIMENSIONALES

Marginal de X

Marginal de Y

Distribución condicionada

Covarianza

Regresión lineal Coeficiente de regresión =

PROBABILIDAD

PROBABILIDAD CONDICIONADA

2 sucesos son independientes si y sólo si

VARIBLAS ALEATORIAS

	DISCRETAS	CONTINUAS
PROBABILIDAD	Función de Probabilidad o Distribución de Probabilidad	Función de Densidad
PROBABILIDAD ACUMULADA	Distribución de Probabilidad Acumulada	Distribución de Probabilidad Acumulada

ESPERANZA MATEMÁTICA O VALOR ESPERADO DE UNA VA

Discretas

Continuas

MOMENTOS (Fórmulas para variables continuas)

Momento de orden r respecto del origen

Momento central de orden r

Momento central de orden 2 o Varianza

COEFICIENTE DE VARIACIÓN

COEFICIENTE DE SIMETRÍA

COEFICIENTE DE CURTOSIS

DISTRIBUCIONES DISCRETAS

BERNOUILLI

BINOMIAL

HIPERGEOMÉTRICA

BINOMIAL MNEGATIVA

GEOMÉTRICA

DE POISSON

DISTRIBUCIONES CONTINUAS

NORMAL

La fórmula es muy complicada. Se utilizan las tablas normalizadas

La Binomial se puede aproximar por la normal.

UNIFORME CONTINUA

Media, mediana y moda coinciden

EXPONENCIAL

<u>TEMA 1.− ESTADÍSTICA DESCRIPTIVA</u>

poblaciones, variables y OBSERVACIONES

Población. Conjunto de seres u objetos acerca de los que se desea obtener información.

Individuo. Cada uno de los miembros de la población.

Variable estadística. Cada una de las características que se miden en los individuos.

Valores. Las diferentes manifestaciones de una variable.

Observación. El conjunto de valores de cada variable medidos en un individuo.

Las variables se suelen clasificar en:

- Cualitativas. Son cualidades de los individuos.
- Ordinales. Los valores pueden ser ordenados.
- Cuantitativas. Miden alguna cualidad cuantificable. Se clasifican en discretas y continuas.

valores de POSICIÓN

Media. Es la suma de todos los valores dividida por el número de valores.

Mediana. Aquel valor para el cual la mitad o menos de la población tienen valores inferiores a él y la mitad o menos de la población tienen valores superiores a él.

Moda. El valor más frecuente de la variable.

la DISPERSIÓN respecto de la media

Varianza. Mide la dispersión o desviación respecto de la media y es igual a la suma de los cuadrados de las desviaciones individuales, dividida por el número de observaciones:

La varianza es igual al promedio de los cuadrados menos el cuadrado de la media:

Desviación típica. Se define como la raíz cuadrada de la varianza

Coeficiente de variación. Es el cociente entre la desviación típica y la media. Se expresa en porcentaje.

DISTRIBUCIONES BIDIMENSIONALES

Las frecuencias se representarán por Fij y fij

Recibe el nombre de **distribución marginal** de X la distribución que tiene la variable X ignorando la variable Y. Y viceversa.

Distribuciones condicionadas:

covarianza

Es una medida de la intensidad de cierta asociación estadística entre dos variables.

También se puede expresar como , que se lee: La covarianza es igual al promedio de los productos cruzados menos el producto de las medias.

2.− REGRESIÓN LINEAL

Es un método para calcular el valor de una variable (Y) a partir de los valores de otras (X):

REGRESIÓN LINEAL simple

A partir de n observaciones (xi yiI se trata de calcular la recta y=a+bx que mejor se ajuste a los datos según el método de los mínimos cuadrados.

En la observación i−ésima:

- El valor previsto de y será:
- El valor observado será yi
- El error será
- El error cuadrático será

La suma de los errores cuadráticos se llama error total:

La recta de regresión de y sobre x es aquella que hace mínimo el error cuadrático total ET. El problema consiste en hallar los valores de a y b que hacen mínimo ET.

La ecuación de la recta de regresión de y sobre x es: A la pendiente de la recta se le llama **coeficiente de regresión.**

Además la recta siempre pasa por el centro de gravedad de la nube de puntos, que es .

Si la covarianza sxy es >0, la recta de regresión es creciente; si sxy es <0, la recta de regresión es decreciente; si es =0, la recta de regresión es cte. (paralela al eje x).

Momentos de la previsión de y ().

La variable prevista tiene la misma media que y:

Varianza de

La covarianza entre e y es igual a la varianza de :

Varianza residual.

Como , se puede demostrar que . A se le llama varianza residual y es la varianza del error residual.

Coeficiente de Determinación.

Se llama Coeficiente de Determinación a la parte de la varianza explicada por la regresión en relación con la varianza de −−>y[Author:FCB].

. Esta última relación nos da una fórmula muy útil a efectos prácticos. Este coeficiente varía entre 0 y 1:

Coeficiente de Correlación.

Se define como la raíz cuadrada positiva del Coeficiente de Determinación.

Varía entre −1 y 1.

Cuanto mayor es (en valor absoluto), mayor es el ajuste de la nube de puntos a la recta, lo que implica que mejores son las estimaciones.

Si =1, la recta es creciente. Si =−1 la recta es decreciente.

PROBABILIDAD

La Probabilidad expresa el valor límite de la frecuencia de un fenómeno. Es un número entre 0 y 1.

espacio muestral y sucesos

Espacio muestral () es el conjunto de los resultados posibles de un experimento aleatorio. Siempre es finito.

Los subconjuntos de se llaman sucesos. es el suceso seguro y es el suceso imposible.

concepto de PROBABILIDAD

El objetivo es atribuir una probabilidad a cada suceso. Se trata de establecer una aplicación P:P()[0,1] que asigne a cada suceso un número entre 0 y 1.

Para que P sea una probabilidad debe cumplir ciertos requisitos:

- Si
- P()=1

PROPIEDADES de la PROBABILIDAD

- Si , entonces
- Si , entonces P(B−A)=P(B)−P(A)
- Si A1 . . . An son disjuntos dos a dos, entonces

ASIGNACIÓN de PROBABILIDADES. regla de laplace

metodos combinatorios

Una permutación de A es una biyección de A en A

Pn=n!

Una variación con repetición de orden r de A es una lista ordenada de r elementos de A.

Una variación con repetición de orden r de A es una lista ordenada de r elementos de A, dónde pueden repetirse elementos.

Una combinación de orden r de A es una lista de r elementos de A

Una combinación con repetición de orden r de A es una lista de r elementos de A, dónde pueden repetirse elementos.

	Importa el orden	No importa el orden
No se permiten repeticiones	VARIACIONES	COMBINACIONES

Se permiten repeticiones	VARIACIONES CON REPETICIÓN	COMBINACIONES CON REPETICIÓN

PROBABILIDAD condicionada

PROPIEDADES de la PROBABILIDAD condicionada

1.− Si A1.......An son disjuntos, entonces

2.−

formula de las PROBABILIDADES totales

Si B1.....Bn son disjuntos dos a dos y

,

entonces

formula de bayes

Si B1.....Bn son disjuntos dos a dos y

,

entonces

sucesos dependientes e independientes

Dos sucesos son independientes si

ESPACIOS PRODUCTO

Este concepto se utiliza cuando un experimento consta de varios experimentos totalmente inconexos entre sí. Entonces se puede describir mediante el producto cartesiano y

independencia de varios sucesos

3 sucesos son independientes si son independientes dos a dos y además

n sucesos son independientes si la probabilidad de la intersección de cualquier número finito de ellos coincide con el producto de las probabilidades:

independencia condicional

Si se verifica , se dice que A y B son independientes condicionalmente a C. Si se sabe que C ha ocurrido, la aparición de B no modifica la probabilidad de A ni al contrario. Los sucesos A y B pueden ser independientes, pero no ser condicionalmente independientes ni a C ni a no C

VARIABLES Aleatorias

Se llama Variable Aleatoria X sobre un espacio probabilístico a una función

variables discretas

El conjunto de pares ordenados (x,p(x)) es una función de probabilidad si cumple:

- P(X=x)=p(x)

Distribución acumulada:

Propiedades de F:

-

variables continuas

Una función f(x) en un intervalo I es una función de probabilidad si cumple:

-

Distribución acumulada:

Propiedades de F:

-

esperanza MATEMÁTICA

Es un concepto análogo a la media aritmética para variables aleatorias.

PROPIEDADES:

- E(c)=c
- E(aX+b)=aE(X)+b
- E(g(X)+h(X))=E(g(X))+E(h(X))

VARIABLES DISCRETAS

Se llama esperanza matemática o valor esperado al número:

Si g(x) es una función de R en R:

VARIABLES CONTINUAS

Se llama esperanza matemática o valor esperado al número:

Si g(x) es una función de R en R:

momentos de una variable aleatoria

Los momentos de una VA son los valores de ciertas funciones de la VA. Generalmente se definen respecto al 0 o respecto al valor esperado.

Momento r−ésimo respecto al origen:

DISCRETA

CONTINUA

Si r=1, entonces

Momento r−ésimo respecto a la media o momento central:

DISCRETA

CONTINUA

Momento central de orden 0:

Momento central de orden 1:

Momento central de orden 2 (Varianza):

Propiedades de la Varianza:

Desviación típica ():

Es la raíz cuadrada de la Varianza.

Coeficiente de variación:

Momento central de orden 3:

Da idea del grado de simetría de la distribución.

Coeficiente de simetría:

Momento central de orden 4:

Da idea del grado de apuntamiento.

Coeficiente de apuntamiento:

otras medidas de CENTRALIZACIÓN y DISPERSIÓN

MEDIANA x es mediana si

DISCRETA

CONTINUA

MODA. Es el valor o valores que hacen máxima la función de probabilidad (DISCRETAS) o la función de densidad (CONTINUAS)

CUANTIL (). Es la solución o soluciones de la ecuación

DISCRETA

CONTINUA

DESVIACIÓN MEDIA. Valor esperado, en valor absoluto, de las desviaciones a la media.

DISCRETA

CONTINUA

<u>MODELOS PROBABILISTICOS DISCRETOS</u>

DISTRIBUCIÓN UNIFORME DISCRETA

X toma valores x1.........xk con probabilidad 1/k. Se dice que X sigue una distribución uniforme discreta de parámetro k.

MEDIA

VARIANZA

DISTRIBUCIÓN de bernouilli

Sólo hay dos resultados posibles: éxito y fracaso. Si p es la probabilidad de éxito, entonces 1−p es la probabilidad de fracaso.

Una VA X que toma 2 valores posibles x1 y x2 con probabilidades p y 1−p sigue una distribución de Bernouilli de parámetro p.

MEDIA

VARIANZA

DISTRIBUCIÓN binomial

Resulta cuando se efectúan varios experimentos de Bernouilli independientes entre sí y queremos medir ewl número k de éxitos entre las n pruebas, teniendo que la probabilidad de éxito en cada prueba es p.

B(n,p):

Distribución de probabilidad:

Media

Varianza

Desviación típica

DISTRIBUCIÓN hipergeometrica

Se aplica en la realización de varios experimentos de Bernouilli, pero que no son independientes entre sí: el resultado de un experimento influye en la probabilidad de éxito de los siguientes.

Tenemos m elementos. Entre esos m elementos hay a que presentan cierta característica que nos interesa. Luego habrá m−a que no la presenten. Extraemos n elementos sin reemplazamiento y queremos saber la probabilidad de que haya k elementos con la característica que nos interesa entre los n extraídos.

Esta VA N sigue una distribución hipergeométrica de parámetros m, a, n.

H(m,a,n):

Distribución de probabilidad:

Media:

Varianza

DISTRIBUCIÓN binomial negativa

En un experimento de Bernouilli, con probabilidad de éxito p, ahora contamos el número n=k+x de repeticiones necesarias para obtener k éxitos.

BN(k,p):

Distribución de probabilidad:

Media:

Varianza:

DISTRIBUCIÓN GEOMÉTRICA

Es el caso particular de la Binomial Negativa cuando k=1. El número de ensayos que hay que realizar para obtener el primer éxito.

G(p):

Distribución de probabilidad:

Media:

Varianza:

DISTRIBUCIÓN de poisson

Sirve para medir la ocurrencia de un determinado suceso a lo largo de un intervalo de tiempo.

La probabilidad de que en un tiempo dt ocurra cierto suceso es . Sea un proceso de Poisson observado durante un tiempo t y con parámetro .

P(t):

Distribución de probabilidad: , con k=0 .... y con

Media

Varianza:

La distribución de Poisson puede utilizarse para aproximar la Distribución Binomial cuando p es pequeña y n grande.

q=1−p

	DISTR PROBABILIDAD	MEDIA	VARIANZA	NOTAS
UNIFORME DISCRETA				
BERNOUILLI				Generalmente éxito=1 y fracaso=0
BINOMIAL				Experimento de Bernouilli repetido n veces de forma independiente
HIPERGEOMETRICA				Experimento de Bernouilli repetido n veces de forma no independiente
BINOMIAL NEGATIVA				Cuantos experimentos k+x hay que hacer para obtener k éxitos
GEOMETRICA				BN con k=1
PIOSSON				Ocurrencia de un determinado suceso a lo

				largo de un periodo de tiempo La Binomial puede aproximarse por la de Poisson

<u>modelos probabilisticos continuos</u>

la DISTRIBUCIÓN normal

Se dice que X sigue una distribución Normal de parámetros , , con b y con si su función de densidad es:

Realmente es la media y es la desviación la típica.

Algunas propiedades importantes son:

- La curva es simétrica respecto a la media. Presenta su máximo en x=y puntos de inflexión en x=
- El área limitada por la curva y el eje x vale 1, lo que garantiza que f(x) es una función de densidad.
- Todos los momentos centrales de orden impar son nulos, por lo tanto el coeficiente de simetría es nulo.
- El momento central de orden cuarto es , por lo tanto el coeficiente de curtosis es nulo.
- La mediana , la moda y la media coinciden

La función de distribución acumulada no puede calcularse analíticamente. Este problema se resuelve mediante la tipificación de la variable.

Definimos una nueva variable Z como , a la cual llamamos variable tipificada. Este cambio de variable nos permite calcular la función de distribución acumulada de una variable normal cualquiera (X) en función de la distribución acumulada de una variable normal de media 0 y varianza 1.

Los valores de se obtienen de la tabla, teniendo en cuenta que ésta incluye sólo los valores positivos (superiores a la media) de . A partir de estos puede calcularse cualquier otra combinación de valores:

APROXIMACIÓN de la binomial por la normal

Sea X una VA que sigue una distribución binomial B(n,p). Si tipificamos la variable restándole la media () y dividiendo por la desviación típica (obtenemos una nueva VA . Entonces, cuando la distribución de la VA Y tiende a la distribución normal estándar N(0,1). Como consecuencia, puede utilizarse la tabla de la normal para calcular de manera aproximada probabilidades binomiales. Hay que tener en cuenta que estamos aproximando una variable discreta por una variable continua.

DISTRIBUCIÓN uniforme continua

Se dice que una VA X sigue una distribución uniforme en el intervalo (a,b) U(a,b) si su función de densidad es:

El valor esperado es:

DISTRIBUCIÓN exponencial

Es un caso particular de la Distribución Gamma.

Se dice que una VA sigue una distribución Gamma de parámetros y (G(,)) si su función de densidad es:

Se dice que una VA X sigue una distribución exponencial de parámetro (E()) si su función de densidad es:

que es la distribución anterior, con =1

La media y la varianza son y

<u>modelos MULTIDIMENSIONALES</u>

DISTRIBUCIONES conjuntas

VARIABLES DISCRETAS

El conjunto de ternas (x,y,p(x,y)) es una distribución de probabilidad conjunta de las VA discretas X e Y si satisface las condiciones:

- p(x,y)0 para todo (x,y)
- P(X=x, Y=y)=p(x,y)

Dadas dos VA discretas X e Y, se denomina función de distribución conjunta F(x,y) a la función:

VARIABLES CONTINUAS

Una función f(x,y) definida en el plano es una función de densidad de probabilidad conjunta de las VA continuas X e Y si satisface las condiciones:

-

La función de distribución conjunta F(x,y) de las VA continuas X e Y cuya función de densidad d eprobablidad es f(x,y) es:

DISTRIBUCIONES MARGINALES

Sean X, Y dos VA discretas con distribución de probabilidad conjunta p(x,y). Las distribuciones marginales de X e Y vienen dadas por:

Sean X, Y dos VA continuas con distribución de probabilidad conjunta f(x,y). Las distribuciones marginales de X e Y vienen dadas por:

momentos de una variable aleatoria BIDIMENSIONAL

VER CASAS

b

a