

TEMA 6

VARIABLES ALEATORIAS UNIDIMENSIONALES

En la sección dedicada a la Estadística Descriptiva hemos estudiado las **variables estadísticas**, estudiándolas como mediciones que se efectúan sobre los individuos de la muestra. Trabajáremos, por tanto con números observados **después** de la realización del experimento.

Si se analizan las variables desde una perspectiva más formal, entendiéndolas como una abstracción **previa** a la realización del experimento, reciben el nombre de **variables aleatorias (v. a.)**, a cuyos posibles resultados se les asocian **probabilidades**, que desempeñan un papel análogo al de las **frecuencias relativas**.

Así, pueden entenderse los desarrollos que se realizarán a continuación como una revisión de los temas de la sección de Estadística Descriptiva, donde los experimentos ya realizados son substituidos por experimentos potenciales, y las frecuencias relativas por probabilidades.

Variable aleatoria.- Matemáticamente, se define una variable aleatoria como una aplicación que a cada suceso elemental del espacio muestral asigna un número real:

Las variables aleatorias se clasifican en **discretas** y **continuas**.

Fig. 6.1 Función de densidad de una variable discreta (a) y continua (b)

Si se considera el experimento que consiste en lanzar dos dados y sumar las puntuaciones de ambos, la X estudiada es una **variable estadística discreta**, siendo los valores del experimento su distribución de frecuencias relativas:

x_i	2	3	4	5	6	7	8	9	10	11	12
f_i	1/20	1/20	2/20	2/20	3/20	4/20	3/20	2/20	1/20	0/20	1/20

Si por el contrario al realizar el experimento se considera lo que puede suceder, analizando los 36 casos posibles y los respectivos casos favorables, obtenemos la **distribución de masa de probabilidad** de la v.a. discreta X :

x_i	2	3	4	5	6	7	8	9	10	11	12
P_i	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Una variable aleatoria es **discreta** si toma un conjunto finito (o infinito numerable) de valores. Estos valores están separados entre sí, y por tanto corresponden a experimentos en los que un mismo resultado puede ocurrir varias veces, y en consecuencia, tiene interés contarlas.

Función de masa de probabilidad.- Sea X una v. a. discreta que toma los valores $x_1, \dots, x_n$, ya ordenados. Se llama **función de masa de probabilidad** a la función que asigna a cada x_i su probabilidad:

$$p_i = P(X = x_i)$$

Su representación gráfica es análoga al diagrama de barras de la Estadística Descriptiva. De la definición se deduce inmediatamente que:

$$0 \leq p_i \leq 1$$

En caso de que el número de valores distintos sea infinito numerable, la segunda propiedad se escribirá: para unos $x_1, \dots, x_n, \dots$ con probabilidades $p_1, \dots, p_n, \dots$

Función de distribución de una v. a..- Es aquella que a cada número real x le asocia la probabilidad de que la variable tome valores menores o iguales a dicho número:

En el caso de que la v.a. X sea discreta, la probabilidad de $F(x) = P(X \leq x)$ se reduce a sumar las probabilidades de los valores menores o iguales a x . Es por tanto una función que acumula la probabilidad hasta x , desempeñando un papel análogo al de las frecuencias relativas acumuladas.

Propiedades de F (caso discreto):

$$0 \leq F(x) \leq 1$$

$$F(-\infty) = 0$$

$$F(+\infty) = 1$$

F creciente.

F es constante entre dos valores consecutivos de la variable.

$$P(a < X < b) = F(b) - F(a).$$

En resumen, F es escalonada, creciente, con saltos en los puntos x_i de amplitud p_i , que asocia a cada valor de la variable la probabilidad acumulada.

MEDIDAS CARACTERÍSTICAS DE UNA VARIABLE ALEATORIA

De modo análogo a la Estadística Descriptiva, en la que definimos unas medidas que caracterizaban a una distribución de frecuencias, podemos definir las mismas medidas para caracterizar a una variable aleatoria discreta (v.a.d.), sin más que reemplazar las frecuencias relativas por probabilidades.

Las más importantes son:

Media.- También llamada **esperanza matemática** o **valor esperado**, y será representada por μ o por $E(X)$. Se calcula con la expresión:

.

Si X fuese infinito numerable, el sumatorio habrá que hacerlo hasta $+\infty$.

Varianza.- Va a denotarse por σ^2 o $Var(X)$. Se definirá como:

Nuevamente hay que advertir que el sumatorio será hasta $+\infty$ para el caso infinito numerable.

Desviación típica.- Se representa por σ o S_n .

Propiedades de la media y la varianza:

$$E(a + bX) = a + b \cdot E(X)$$

$$Var(a + bX) = b^2 \cdot Var(X)$$

$$Var(X) = E(X^2) - [E(X)]^2$$

Si g es una función, $g(X)$ es una v. a. que cumple:

- Si X e Y son v. a. cualesquiera, se cumple $E(X + Y) = E(X) + E(Y)$
- Si X e Y son v. a. independientes, se cumple $E(X \cdot Y) = E(X) \cdot E(Y)$
- Si X e Y son v. a. independientes, se cumple $Var(X + Y) = Var(X) + Var(Y)$
- Se cumple la desigualdad de Tchebyshev.

Tipificación de una v. a.- Al igual que se vio en la Estadística Descriptiva, las v.a. pueden tipificarse.

Se conserva la propiedad de que su media es cero y su desviación típica es uno.

PRINCIPALES DISTRIBUCIONES DISCRETAS

I.- Distribución uniforme discreta.- Una v.a. X se dice que tiene una distribución uniforme sobre n puntos $(x_1, \dots, x_n)$ si todos los puntos tienen la misma probabilidad.

Función de masa de probabilidad:

Media:

Varianza:

II.- Distribución binomial.- Hay pruebas en las que interesa distinguir entre los resultados, denominados “Éxito” y “fracaso”, con probabilidades p y q respectivamente, tales que $p + q = 1$. Si el experimento consiste en realizar varias veces la **misma prueba** (la probabilidad de Éxito es igual en cada nueva prueba) y además, las pruebas se realizan con **independencia**, diremos que estamos ante un proceso de Bernoulli.

Si en un proceso de Bernoulli se considera la v. a. $X = \text{“nº de Éxitos en } n \text{ pruebas”}$, se dice que esta variable aleatoria sigue una **distribución binomial** de parámetros n y p , y se denota por la expresión $X \sim Bi(n, p)$. El número de Éxitos está representado por k .

Función de masa de probabilidad:

Media:

$$E(X) = n \cdot p$$

Varianza:

$$Var(X) = n \cdot p \cdot q$$

EJEMPLO (Binomial):

Una sala de ordenadores tiene 10 terminales. La probabilidad de que un terminal esté ocupado es de 0.9. Calcular la probabilidad de que 7 más terminales estén ocupados.

Se define $X = \text{"n}^\circ \text{ de terminales ocupados"}$, y se considera que el suceso A es que un terminal está ocupado. Por tanto, $P(A) = 0.9$. Por tanto $P(\text{"terminal libre"}) = 0.1$.

$X \sim \text{Bi}(10, 0.9)$. La probabilidad de que siete o más terminales estén ocupados es igual a la suma de la probabilidad de que 7 estén ocupados, más la probabilidad de que estén ocupados 8, más la probabilidad de que estén ocupados 9, más la probabilidad de que estén ocupados 10.

La probabilidad de que los terminales estén ocupados es de la forma:

$$P(X \geq 7) = P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10) = 0.0574 + 0.1937 + 0.3874 + 0.3487 = 0.9872$$

III.- Distribución de Poisson.-La generalización de un suceso de Bernoulli a un soporte continuo se llama **proceso de Poisson**, al que se exigen dos condiciones: que sea estable en el tiempo y que los sucesos llamados "éxito" aparezcan independientemente. En estas condiciones, la v.a. X ; "nº de éxitos ocurridos en el intervalo considerado" sigue una distribución de Poisson de parámetro $\lambda = np$, y se denotará de la forma $X \sim \text{Pois}(\lambda)$.

Función de masa de probabilidad:

con $k=0, 1, 2, \dots$

Media:

$$E(X) = \lambda$$

Varianza:

$$\text{Var}(X) = \lambda$$

Propiedad.-

EJEMPLO (Poisson):

La probabilidad de que un programa en PASCAL correctamente escrito aborte por causas desconocidas es de 0.002. Se ejecutan 1200 programas correctamente escritos. Calcular la probabilidad de que aborten a lo sumo 5.

Se define $X = \text{"n}^\circ \text{ de programas que abortan"}$, y se considera que el suceso A es que un programa aborte. Por tanto, $P(A) = 0.002$. Por tanto $P(\text{"no abortar"}) = 1 - P(A) = 0.998$.

Se podría pensar en resolverlo mediante la binomial, pero como se da el caso de que $n \rightarrow \infty$ y $p \rightarrow 0$, entonces sabemos que $X \sim \text{Pois}(2.4)$. ($0.002 \cdot 1200 = 2.4$).

Lo que se pide es la $P(X \leq 5)$. Aplicando la fórmula nos queda que:

$$P(X=k) = e^{-\lambda} \cdot (\lambda^k / k!)$$

$$P(X \leq 5) = P(X = 0) + \dots + P(X = 5) = 0.097 + 0.2177 + 0.2613 + 0.209 + 0.1254 + 0.0602 = 0.9643$$

III.- Distribución Hipergeométrica.- Se considera una población finita de N objetos de los cuales k son de la clase S y los restantes $N-k$ son de la clase M . Se toma una muestra aleatoria, sin

reemplazamiento, de tamaño n y definimos X como el “número de objetos de la clase S entre los n extraídos”. Si la muestra se hubiese tomado con reemplazamiento, entonces X seguiría a una binomial (n, p) , con p , pero al no haber reemplazamiento, las sucesivas extracciones son dependientes, y aunque la probabilidad de éxito incondicional es la misma, la probabilidad de éxito condicionada a lo que ya ha sucedido varía de prueba en prueba. En estas condiciones se dice que X sigue una distribución hipergeométrica. Se denota por $X \sim H(n, N, p)$

Para ver esto de una forma más sencilla, se puede decir que N es el tamaño del conjunto del cual se va a extraer la muestra, n es el tamaño de la muestra, y p es la probabilidad de éxito dentro de la muestra.

Función de masa de probabilidad:

r es el número de elementos de N que cumplen la condición requerida por el problema, y $N-r$ es el número de elementos que no la cumplen.

Media:

$$E(X) = np$$

Varianza:

$$Var(X) =$$

La distribución hipergeométrica se utiliza en el muestreo de poblaciones finitas sin reemplazamiento.

EJEMPLO (D. Hipergeométrica):

Se tiene una baraja de 40 cartas y se extraen 5 sin reemplazamiento. Probabilidad de que 3 seanoros.

Se define X = “númerooros en las cinco cartas extraídas”.

$$\text{Se ve que } X \sim H(40, 10, 5/40) = H(40, 10, 0.25)$$

$$r = 10 \text{ (oros en la baraja)}$$

IV.- Distribución geométrica (o de Pascal).- En las mismas condiciones del experimento binomial, consideramos la v.a. X como el “número de fracasos antes del primer éxito”. La variable así definida sigue una distribución geométrica de parámetro p , y se denota $X \sim \text{Pascal}(p)$.

Función de masa de probabilidad:

Media:

Varianza:

EJEMPLO (Pascal):

Una pareja desea tener hijos hasta la primera niña. Se pide la probabilidad de que tengan más de cuatro niños.

Existen dos formas de realizar este ejercicio. Una es sumar las probabilidades desde 5 niños hasta infinitos niños, y la otra es calcular la probabilidad de que tengan 4 o menos hijos, aprovechando que $P(X \leq 4) = 1 - P(X > 4)$. Por tanto:

$$P(X > 4) = 1 - (P(X = 4) + P(X = 3) + P(X = 2) + P(X = 1) + P(X = 0)).$$

Calculamos aplicando la fórmula:

estamos considerando que la probabilidad de que nazca niño o niña es igual, y que su valor es 0.5. Luego $P(X > 4) = 1 - (0.0313 + 0.0625 + 0.1250 + 0.25 + 0.5) = 0.03125$

V.- Distribución binomial negativa.- Es una generalización de la distribución anterior. La v.a. considerada es X ; “número de fracasos antes del éxito n ”, que sigue una distribución binomial negativa con parámetros “ n ” y “ p ”. Se denotará como $X \sim BN(j, p)$.

Función de masa de probabilidad:

Media:

Varianza:

RELACIONES ENTRE LAS DISTRIBUCIONES

Sea X una distribución $H(N, n, r)$. En el caso de que , se puede aproximar por una $Bi(n, p)$.

Sea $X \sim Bi(n, p)$. Si se cumple que $n \cdot p < 5$ y $p < 0.1$, entonces la distribución de X se puede aproximar por una $Pois(\hat{\lambda})$, donde $\hat{\lambda} = n \cdot p$.

Sea $X \sim BN(j, p)$. Con $n = j$ se puede aproximar por una Pascal de parámetro p'

VARIABLES ALEATORIAS CONTINUAS

Una variable aleatoria es continua si puede tomar cualquier valor en un intervalo. Su función de distribución se define:

Propiedades de F (caso continuo):

$$0 \leq F(x) \leq 1$$

$$F(-\infty) = 0$$

$$F(\infty) = 1$$

F creciente.

F es continua

$$P(a < X < b) = F(b) - F(a).$$

En la práctica no es posible conocer el valor exacto de una realización de una v.a. continua, ya que al medir lo que se hace es clasificarla dentro de un intervalo más o menos amplio, dependiendo de la precisión del aparato de medida. En este principio se basa la construcción del histograma. Si se hace que la

amplitud de los intervalos de clase sea cada vez menor, puede verse que los histogramas tienen una curva cada vez más suave. El límite de los histogramas así construidos, cuando $n \rightarrow \infty$, y la amplitud de los intervalos tiende a cero, es una curva que recibe el nombre de **función de densidad** de la v.a., y que se nota matemáticamente $f(x) = F'(x)$.

Propiedades de f . Para que una función sea función de distribución, **debe** cumplir las siguientes condiciones:

$$f(x) \geq 0$$

Hay que comentar que

Relación entre F y f .

$$f(x) = F'(x)$$

Fig. 6.2.- La probabilidad es el área bajo la función de densidad sobre el conjunto de interés

Puede observarse que $P(X = x) = 0$, para todo $x \in \mathbb{R}$.

Desigualdad de Tchebyshev. Si X es v.a. con media μ y desviación típica σ , y $k > 0$, entonces

PRINCIPALES DISTRIBUCIONES CONTINUAS

I.- Distribución uniforme. Una v.a. que toma valores en un intervalo con probabilidad constante es una distribución uniforme ($X \sim U(a, b)$).

Función de densidad:

Media:

Varianza:

II.- Distribución exponencial. La distribución exponencial surge cuando, en un proceso de Poisson, estamos interesados en la v.a. X ; “**tiempo transcurrido entre dos éxitos consecutivos**”. Se dice entonces que esta v.a. sigue una distribución exponencial con parámetro λ ($X \sim \text{Exp}(\lambda)$), donde λ es el número de éxitos por unidad de tiempo. X puede tomar cualquier valor entre cero e infinito.

Función de densidad:

Para resolver una probabilidad, se sabe que:

Media:

Varianza:

EJEMPLO:

Supóngase que la duración (en horas) de cierto tubo de radio es una variable aleatoria continua X con función de densidad .

a) ¿Cuál es la probabilidad de que un tubo dure menos de 200 horas si todavía funciona tras 150

horas de servicio?

Se aplica la fórmula sabiendo que el Valor a es 150 y que el Valor b es 200. Luego, la respuesta será;
 $(e^{-0.01 \cdot 200}) - (e^{-0.01 \cdot 150}) = 0.0878$

b) ¿Cuál es la probabilidad de que si se instalan tres de estos tubos en un conjunto, solamente uno de ellos no se haya estropeado al cabo de 150 horas?

Lo primero será hallar la probabilidad de que un tubo cualquiera funcione al cabo de esas 120 horas. Para ello se hace la integral entre 150 horas y cero, que es igual al valor de la probabilidad de que se produzca el error a las 150 horas $(e^{-0.01 \cdot 150}) - (e^{-0.01 \cdot 0}) = 0.7769$

Ahora se puede resolver aplicando una $Bi(3, 0.7769)$:

0.1160

III.- Distribución normal.- Es la más importante de las distribuciones. También es conocida como Distribución de Gauss o Distribución de los errores, pues modeliza multitud de fenómenos que se dan en la naturaleza, como por ejemplo los errores en una medición. Se nota como $X \sim N(\mu, \sigma)$. Una variable aleatoria continua X sigue una distribución normal de parámetros μ y σ , siendo $\mu \in \mathbb{R}$ y $\sigma > 0$, si su **función de densidad** es de la forma:

R

Media:

$$E(X) = \mu$$

Varianza:

$$Var(X) = \sigma^2$$

En la práctica, si $n > 30$ y $npq > 5$, una binomial $Bi(n, p)$ se aproxima por una normal de la forma. Otra aproximación útil, esta referente a la distribución de Poisson es que cuando $\lambda > 5$ puede aproximarse por una .

La normal tipificada o estandarizada se calcula:

EJEMPLO:

Cierto tipo de cuadernos escolares son empaquetados en mazos de 25 cuadernos cada uno por una máquina que realiza el recuento y empaqueta de forma automática. Con el fin de verificar la exactitud de la máquina, los mazos se pesan antes de enviarlos a las papelerías. Se sabe que el peso de los cuadernos individuales es una variable aleatoria con distribución normal de media 40 g. y desviación típica 3 g. Si se considera que un mazo tiene 25 cuadernos cuando su peso está comprendido entre 975 y 1025 gramos, hallar la probabilidad de que un mazo que tiene 24 cuadernos sea considerado como si tuviese 25.

Se sabe que un mazo tiene 25 cuadernos, y que el peso de un cuaderno sigue una $N(40, 3)$. Se considera que un mazo tiene 25 cuadernos si su peso está en (975, 1025). Entonces la probabilidad de que un mazo de 24 sea considerado como uno de 25 es igual a la probabilidad de que el peso de ese mazo esté en (975, 1025).

Se halla la esperanza de peso, y la varianza de un mazo de 24:

$$E[S_{24}] = n \hat{\mu} = \text{cuadernos} \hat{\mu} = \text{peso de los cuadernos} = 24 \hat{\mu} = 40 = 960 \text{ gramos.}$$

$$\text{Var}[S_{24}] =$$

Para poder usar la tabla de la normal, hay que tipificar los datos. Lo hay que calcular es:

$$P(975 \hat{\mu} \leq S_{24} \hat{\mu} \leq 1025)$$

Por tanto se tipifica aplicando , y se realiza el cálculo:

$$= 0.9999 - 0.8461 = 0.1539$$

Los valores $F(4.42)$ y $F(1.02)$ se buscan en las tablas de la normal tipificada

Propiedades de la Distribución normal:

La suma de v.a. normales independientes sigue siendo una v.a. normal:

- **Teorema central del Límite.-** Si sumamos muchas v.a. independientes, la suma se puede aproximar por una v.a. normal.

Sea una sucesión de v.a. independientes con la misma distribución, de media μ , y varianza σ^2 . Entonces se verifica que:

Estadística 1º E.T.I.S. Facultad de Informática da Coruña Curso 1.997-1.998