Joseline Jamileth Montiel Vazquez

Nombre del profesor:

Distribución de variables discretas

Estadística

1er Cuatrimestre

Escolarizado

Pichucalco, Chiapas a 04 de diciembre de 2020.

TIPOS DE VARIABLES:

- Variables discretas:

Las variables discretas son variables numéricas que tienen un número contable de valores entre dos valores cualesquiera. Una variable discreta siempre es numérica. Por ejemplo, el número de quejas de los clientes o el número de fallas o defectos.

- Variables continuas:

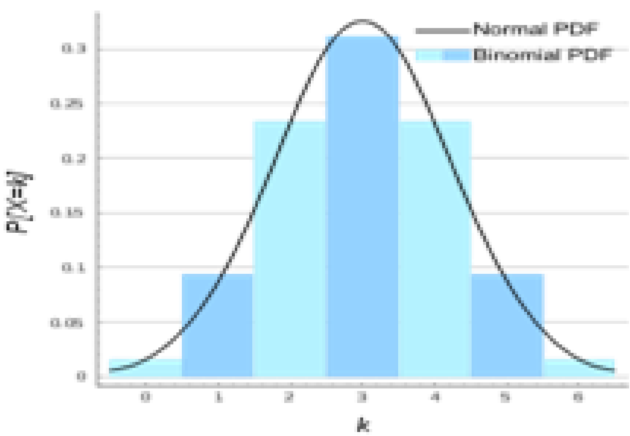
Las variables continuas son variables numéricas que tienen un número infinito de valores entre dos valores cualesquiera. Una variable continua puede ser numérica o de fecha/hora. Por ejemplo, la longitud de una pieza o la fecha y hora en que se recibe un pago.

DISTRIBUCION DE VARIABLES DISCRETAS:

- Distribución binomial:

En estadística, la distribución binomial es una distribución de probabilidad discreta que mide el número de éxitos en una secuencia de n ensayos independientes de Bernoulli con una probabilidad fija p de ocurrencia del éxito entre los ensayos.

Un experimento de Bernoulli se caracteriza por ser dicotómico, esto es, sólo son posibles dos resultados. A uno de estos se denomina éxito y tiene una probabilidad de ocurrencia p y al otro, fracaso, con una probabilidad q = 1 – p. En la distribución binomial el anterior experimento se repite n veces, de forma independiente, y se trata de calcular la probabilidad de un determinado número de éxitos. Para n = 1, la binomial se convierte, de hecho, en una distribución de Bernoulli.



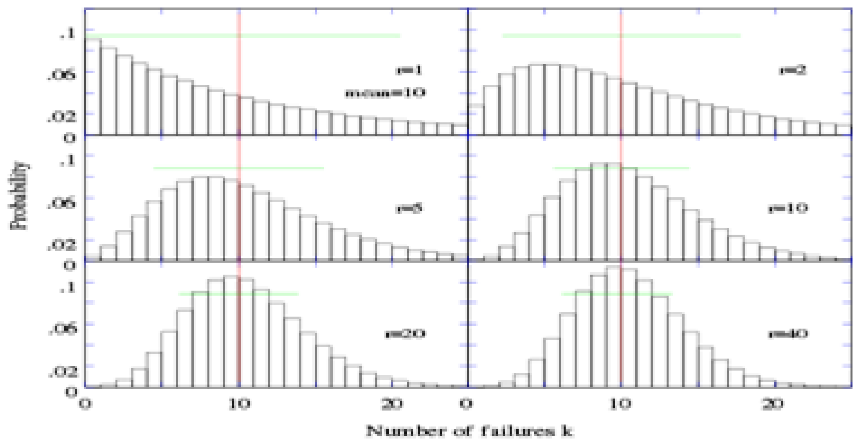
Para representar que una variable aleatoria X sigue una distribución binomial de parámetros n y p, se escribe:

- Distribución binomial negativa:

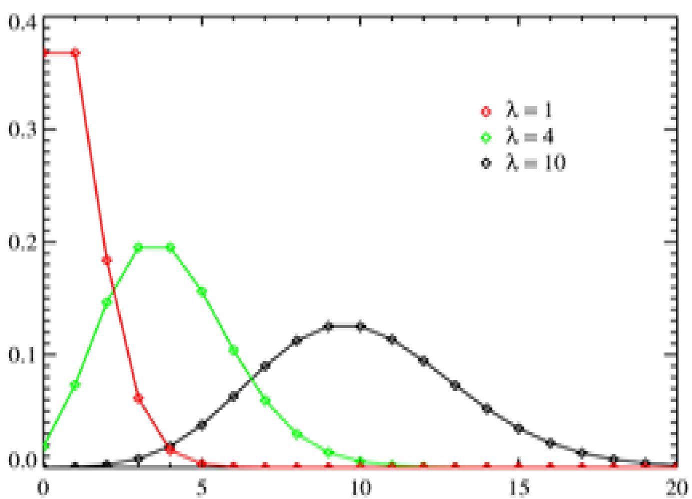
En estadística la distribución binomial negativa es una distribución de probabilidad discreta que incluye a la distribución de Pascal.

El número de experimentos de Bernoulli de parámetro θ independientes realizados hasta la consecución del k-ésimo éxito es una variable aleatoria que tiene una distribución binomial negativa con parámetros k y θ.

La distribución geométrica es el caso concreto de la binomial negativa cuando k = 1.



- Distribución de poisson:

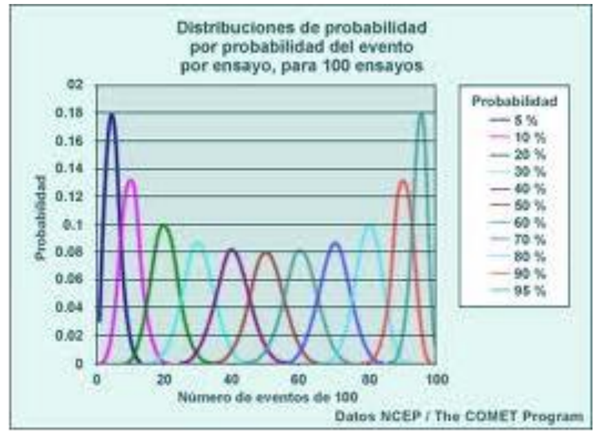


En teoría de probabilidad y estadística, la distribución de Poisson es una distribución de probabilidad discreta. asi tiempo fijo si estos eventos ocurren con una frecuencia media conocida y son independientes del tiempo discurrido desde el último evento.

Fue descubierta por Siméon-Denis Poisson, que la dio a conocer en 1838 en su trabajo Recherches sur la probabilité des jugements en matières criminelles et matière civile (Investigación sobre la probabilidad de los juicios en materias criminales y civiles).

- Distribución geométrica:

En teoría de probabilidad y estadística, la distribución geométrica es cualquiera de las dos distribuciones de probabilidad discretas siguientes:



- la distribución de probabilidad del número X del ensayo de Bernoulli necesaria para obtener un éxito, contenido en el conjunto { 1, 2, 3,…} o
- la distribución de probabilidad del número Y = X − 1 de fallos antes del primer éxito, contenido en el conjunto { 0, 1, 2, 3,… }.

Cuál de éstas es la que uno llama “la” distribución geométrica, es una cuestión de convención y conveniencia.

- Distribución hipergeométrica:

En teoría de la probabilidad la distribución hipergeométrica es una distribución discreta relacionada con muestreos aleatorios y sin reemplazo. Supóngase que se tiene una población de N elementos de los cuales, d pertenecen a la categoríaA y N-d a la B. La distribución hipergeométrica mide la probabilidad de obtener x ( ) elementos de la categoríaA en una muestra de n elementos de la población original

- Distribución de Bernoulli:

En teoría de probabilidad y estadística, la distribución de Bernoulli (o distribución dicotómica), nombrada así por el matemático ycientífico suizo Jakob Bernoulli, es una distribución de probabilidad discreta, que toma valor 1 para la probabilidad de éxito (p) y valor 0 para la probabilidad de fracaso (q = 1 − p).

Si X es una variable aleatoria que mide “número de éxitos”, y se realiza un único experimento con dos posibles resultados (éxito o fracaso), se dice que la variable aleatoria X se distribuye como una Bernoulli de parámetro p.

La fórmula será:

f(x) = px(1 − p)1 − x con x = {0,1}

Su función de probabilidad viene definida por:

Un experimento al cual se aplica la distribución de Bernoulli se conoce como Ensayo de Bernoulli o simplemente ensayo, y la serie de esos experimentos como ensayos repetidos.

DISTRIBUCION DE VARIABLES DISCRETAS:

Una distribución continua describe las probabilidades de los posibles valores de una variable aleatoria continua. Una variable aleatoria continua es una variable aleatoria con un conjunto de valores posibles (conocido como el rango) que es infinito y no se puede contar.

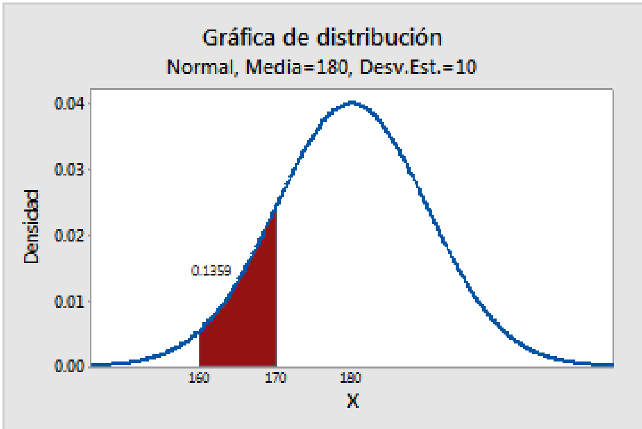
Las probabilidades de las variables aleatorias continuas (X) se definen como el área por debajo de la curva de su PDF. Por lo tanto, solo los rangos de valores pueden tener una probabilidad diferente de cero. La probabilidad de que una variable aleatoria continua equivalga a algún valor siempre es cero.

- Distribución de pesos:

La distribución normal continua puede describir la distribución del peso de hombres adultos. Por ejemplo, usted puede calcular la probabilidad de que un hombre pese entre 160 y 170 libras.

Gráfica de distribución del peso de hombres adultos:

El área sombreada debajo de la curva en este ejemplo representa el rango de 160 a 170 libras. El área de este rango es 0.136; por lo tanto, la probabilidad de que un hombre seleccionado aleatoriamente pese entre 160 y 170 libras es de 13.6%. Toda el área por debajo de la curva equivale a 1.0.



Sin embargo, la probabilidad de que X sea exactamente igual a algún valor siempre es cero, porque el área por debajo de la curva en un punto individual, que no tiene anchura, es cero. Por ejemplo, la probabilidad de que un hombre pese exactamente 190 libras es cero. Podría calcular una probabilidad diferente de cero de que un hombre pese más de 190 libras, menos de 190 libras o entre 189.9 y 190.1 libras, pero la probabilidad de que pese exactamente 190 libras es cero.

METODOS DE MUESTREO:

- Probabilistico:

Es requisito que todos y c/u de los elementos de la población tengan la misma probabilidad de ser seleccionados (azar). Se debe tener disponible un listado completo de todos los elementos de la población, a esto se le llama MARCO DE MUESTREO.

- Aleatorio simple:

Cada sujeto tiene una probabilidad igual de ser seleccionado para el estudio. Se necesita una lista numerada de las unidades de la población que se quiere muestrear.

Opciones:

◦ Fichas de lotería o bolitas numeradas

◦ Tabla de números aleatorios

- Estratificado:

Cuando la muestra incluye subgrupos representativos (estratos) de los elementos de estudio con características específicas: urbano, rural, nivel de instrucción, año académico, carrera, sexo, grupo étnico, edad, paridad etc. En cada estrato para obtener el tamaño de la muestra se puede utilizar el muestreo aleatorio o sistemático.

INTERVALOS DE CONFIANZA:

Un intervalo de confianza es una técnica de estimación utilizada en inferencia estadística que permite acotar un par o varios pares de valores, dentro de los cuales se encontrará la estimación puntual buscada (con una determinada probabilidad).

Un intervalo de confianza nos va a permitir calcular dos valores alrededor de una media muestral (uno superior y otro inferior). Estos valores van a acotar un rango dentro del cual, con una determinada probabilidad, se va a localizar el parámetro poblacional.

Intervalo de confianza = media +- margen de error

Conocer el verdadero poblacional, por lo general, suele ser algo muy complicado. Pensemos en una población de 4 millones de personas. ¿Podríamos saber el gasto medio en consumo por hogar de esa población? En principio sí. Simplemente tendríamos que hacer una encuesta entre todos los hogares y calcular la media. Sin embargo, seguir ese proceso sería tremendamente laborioso y complicaría bastante el estudio.

Ante situaciones así, se hace más factible seleccionar una muestra estadística. Por ejemplo, 500 personas. Y sobre dicha muestra, calcular la media. Aunque seguiríamos sin saber el verdadero valor poblacional, podríamos suponer que este se va a situar cerca del valor muestral. A esa media le sumamos el margen de error y tenemos un valor del intervalo de confianza. Por otro lado, le restamos a la media ese margen de error y tendremos otro valor. Entre esos dos valores estará la media poblacional.

En conclusión, el intervalo de confianza no sirve para dar una estimación puntual del parámetro poblacional, si nos va a servir para hacernos una idea aproximada de cuál podría ser el verdadero de este. Nos permite acotar entre dos valores en dónde se encontrará la media de la población.