



KARINA MENDOZA CRUZ
LIC. RODRIGO MANUEL BRAVO LOPEZ
ANTOLOGIA
08/11/22
BIOESTADISTICA

Modelos de distribución de probabilidad

MODELOS DISCRETOS

Los modelos discretos, son modelos de probabilidad de variable aleatoria discreta. Los más importantes son los modelos de BERNOULLI (especialmente "la distribución binomial") y la "distribución de Poisson".

Distribución Binomial. El campo de variación de la variable es $\{0,1,2, 3,\dots, n\}$ y la función de cuantía es: para valores de $x= 0,1,2, \dots,n$ siendo $n \in \mathbb{N}$, $p \in [0,1]$ y $q=1-p$

Distribución de Poisson

Formalmente: dada una variable aleatoria X con campo de variación $X \in \{0,1, 2, \dots, \infty\}$, es decir $X \in \mathbb{N}$ cuya función de cuantía sea: siendo λ un parámetro positivo diremos que X sigue una distribución de Poisson de parámetro λ , $X \sim P(\lambda)$. Situaciones que modeliza: Se observa la ocurrencia de hechos de cierto tipo durante un período de tiempo o a lo largo de un espacio, considerados unitarios El tiempo (o el espacio) pueden considerarse homogéneos, respecto al tipo de hechos estudiados, al menos durante el período experimental; es decir, que no hay razones para suponer que en ciertos momentos los hechos sean más probables que otros.

Distribución Hipergeométrica Dada la siguiente situación: Una población constituida por N individuos en total. De los cuales N_p individuos son del tipo A , y N_q individuos son del tipo \bar{A} .

MODELOS CONTINUOS Distribución Uniforme (de V. Continua)

Dada una variable aleatoria continua, X , definida en el intervalo $[a, b]$ de la recta real, diremos que X tiene una distribución uniforme en el intervalo $[a, b]$ cuando su función de densidad sea: $X \sim U([a, b])$.

Distribución Exponencial

Dada una variable aleatoria continua, X , definida para valores reales positivos. diremos que X tiene una distribución exponencial de parámetro a cuando su función de densidad sea: $f(x) = a e^{-a x}$ para $x \geq 0$ (siendo el parámetro a positivo).

Distribución Normal

La distribución normal es la más importante de todas las distribuciones de probabilidad. Es una distribución de variable continua con campo de variación $[-\infty, \infty]$, que queda especificada a través de dos parámetros (que acaban siendo la media y la desviación típica de la distribución).

Importancia de la distribución Normal.

a) Enorme número de fenómenos que puede modelizar: Casi todas las características cuantitativas de las poblaciones muy grandes tienden a aproximar su distribución a una distribución normal.

b) Muchas de las demás distribuciones de uso frecuente, tienden a distribuirse según una Normal, bajo ciertas condiciones.

Distribuciones Binomial y Poisson.

Distribución Binomial

Una distribución binomial es una distribución de probabilidad discreta que describe el número de éxitos al realizar n experimentos independientes entre sí, acerca de una variable aleatoria.

Propiedades de la distribución binomial

Para que una variable aleatoria se considere que sigue una distribución binomial, tiene que cumplir las siguientes propiedades:

En cada ensayo, experimento o prueba solo son posibles dos resultados (éxito o fracaso).

La probabilidad del éxito ha de ser constante. Esta se representa mediante la letra p . La probabilidad de que salga cara al lanzar una moneda es 0,5 y esta es constante dado que la moneda no cambia en cada experimento y las probabilidades de sacar cara son constantes

La probabilidad de fracaso ha de ser también constante. Esta se representa mediante la letra $q = 1 - p$. Es importante fijarse que, mediante esa ecuación, sabiendo p o sabiendo q , podemos obtener la que nos falte.

Distribución de Poisson

La Distribución de Poisson se llama así en honor a Simeón Dennis Poisson (1781-1840), francés que desarrolló esta distribución basándose en estudios efectuados en la última parte de su vida.

La distribución de Poisson es una distribución de probabilidad discreta que se aplica a las ocurrencias de algún suceso durante un intervalo determinado.

Distribución normal.

La distribución normal, distribución de Gauss o distribución gaussiana, es la distribución de probabilidad individual más importante.

La distribución normal estándar

La distribución normal estándar, es aquella distribución normal que tiene una media igual a cero, y una desviación estándar igual a uno.

Otras distribuciones discretas y continuas

Distribución Hipergeométrica

La distribución Hipergeométrica es especialmente útil en todos aquellos casos en los que se extraigan muestras o se realicen experiencias repetidas sin devolución del elemento extraído o sin retornar a la situación experimental inicial.

La distribución Gamma

Este modelo es una generalización del modelo Exponencial ya que, en ocasiones, se utiliza para modelar variables que describen el tiempo hasta que se produce p veces un determinado suceso.

La distribución de Cauchy

Se trata de un modelo continuo cuya integral nos proporciona la función de distribución. Propiedades de la distribución de Cauchy Se trata de un ejemplo de variable aleatoria que carece de esperanza (y, por tanto, también de varianza o cualquier otro momento).

Muestreo aleatorio simple.

El muestreo aleatorio simple es un subconjunto de una muestra elegida de una población más grande. Cada individuo se elige al azar y por pura casualidad.

Uso de números aleatorios

El método de uso de números aleatorios es un método alternativo que también implica la numeración de la población.

Justificación del muestreo.

En vez de tomar un censo completo, los procedimientos de muestreo estadístico se han convertido en la herramienta preferida en la mayoría de las situaciones de investigación. Existen tres razones principales para extraer una muestra.

Estadísticos muestrales.

Distribuciones. En estadística un estadístico (muestral) es una medida cuantitativa, derivada de un conjunto de datos de una muestra, con el objetivo de estimar o inferir características de una población o modelo estadístico.

Propiedades de los estimadores

Sesgo: Se denomina sesgo de un estimador a la diferencia entre la esperanza (o valor esperado) del estimador y el verdadero valor del parámetro a estimar.

Eficiencia: Un estimador es más eficiente o preciso que otro, si la varianza del primero es menor que la del segundo.

Convergencia: Para estudiar las características de un estimador no solo basta con saber el sesgo y la varianza, sino que además es útil hacer un análisis de su comportamiento y estabilidad en el largo plazo, esto es, su comportamiento asintótico.

Consistencia: También llamada robustez, se utilizan cuando no es posible emplear estimadores de mínima varianza, el requisito mínimo deseable para un estimador es que a medida que el tamaño de la muestra crece, el valor del estimador tiende a ser el valor del parámetro, propiedad que se denomina consistencia.

Obtención de estimadores.

Método por Analogía.

Consiste en aplicar la misma expresión formal del parámetro poblacional a la muestra, generalmente, estos estimadores son de cómoda operatividad, pero en ocasiones presentan sesgos y no resultan eficientes.

Método de los momentos.

Consiste en tomar como estimadores de los momentos de la población a los momentos de la muestra.

Estimadores máximo - verosímiles.

La verosimilitud consiste en otorgar a un estimador/estimación una determinada "credibilidad" una mayor apariencia de ser el cierto valor(estimación) o el cierto camino para conseguirlo(estimador).

Estimación por intervalos de confianza

La estimación por intervalos consiste en establecer el intervalo de valores donde es más probable se encuentre el parámetro.

Contraste de hipótesis.

Una hipótesis estadística es una asunción relativa a una o varias poblaciones, que puede ser cierta o no.

Construcción de Test de hipótesis.

Seis pasos básicos para configurar y realizar correctamente una prueba de hipótesis.

1. Especificar las hipótesis.
2. Elegir un nivel de significancia (también denominado alfa o α).
3. Determinar la potencia y el tamaño de la muestra para la prueba.
4. Recolectar los datos.
5. Comparar el valor p de la prueba con el nivel de significancia.
6. Decidir si rechazar o no rechazar la hipótesis nula.

Contraste de hipótesis paramétricas.

Es la técnica estadística que se usa para estudiar si una determinada afirmación acerca de cierto parámetro poblacional es confirmada o invalidada por los datos de una muestra extraída de dicha población.

Test para poblaciones normales

Una población

En la inferencia sobre una variable numérica en una población, el objetivo principal de los test de hipótesis es contrastar el valor de alguna medida de posición (media o mediana), de dispersión (varianza) o de algún otro parámetro poblacional.

Test para poblaciones binomiales y de Poisson

Nos encontramos con un modelo derivado de un proceso experimental puro, en el que se plantean las siguientes circunstancias. Se realiza un número n de pruebas (separadas o separables).

Test basado en el estadístico Chi-cuadrado

Esta prueba puede utilizarse incluso con datos medibles en una escala nominal. La hipótesis nula de la prueba Chi-cuadrado postula una distribución de probabilidad totalmente especificada como el modelo matemático de la población que ha generado la muestra.

Test de bondad de ajuste.

La bondad de ajuste de un modelo estadístico describe lo bien que se ajusta un conjunto de observaciones.

Prueba de bondad de ajuste chi cuadrado χ^2

El procedimiento de la prueba requiere una muestra aleatoria de tamaño n proveniente de la población cuya distribución de probabilidad es desconocida.

Test de heterogeneidad

La heterogeneidad estadística es la presencia de diferencias entre los efectos calculados de la intervención, que son mayores que lo que es de esperar si se debieran solamente a las variaciones al azar (muestrales).

Test de homogeneidad

Se plantea el problema de la existencia de homogeneidad entre r poblaciones, para lo cual se realizan muestras independientes en cada una de ellas.

Test de homogeneidad

Se plantea el problema de la existencia de homogeneidad entre r poblaciones, para lo cual se realizan muestras independientes en cada una de ellas.

Tablas de Contingencia

. Una tabla de contingencia es una herramienta utilizada en la rama de la estadística, la cual consiste en crear al menos dos filas y dos columnas para representar datos categóricos en términos de conteos de frecuencia.