**DISTRIBUCIONES DE PROBABILIDAD CON VARIABLES ALEATORIAS CONTINUAS**

Se llama distribución de probabilidad al conjunto de una variable aleatoria y de las probabilidades asociadas a cada valor que pueda tomar. La distribución de probabilidad de una variable aleatoria proporciona una probabilidad para cada valor posible, y estas probabilidades deben sumar 1. Las distribuciones de probabilidad se clasifican como continuas y discretas. En la distribución de probabilidad discreta está permitido tomar sólo un número limitado de valores.

En una distribución de probabilidad continua, la variable que se está considerando puede tomar cualquier valor dentro de un intervalo dado. Las distribuciones continuas son una forma conveniente de presentar distribuciones discretas que tienen muchos resultados posibles, todos muy cercanos entre sí.

**Valor esperado de una variable aleatoria**

El valor esperado es una idea fundamental en el estudio de las distribuciones de probabilidad. Es un promedio ponderado de los resultados que se esperan en el futuro.

Para obtener el valor esperado de una variable aleatoria discreta, se multiplica cada valor que la variable pueda tomar por la probabilidad de presentación de ese valor y luego se suman esos productos. El valor esperado pesa cada resultado posible con respecto a la frecuencia con que se espera que se presente. En consecuencia, las presentaciones más comunes tienen asignadas un peso mayor que las menos comunes.

El valor esperado también puede ser obtenido a partir de estimaciones subjetivas. En ese caso, el valor esperado no es más que la representación de las convicciones personales acerca del resultado posible. En muchas situaciones, encontraremos que es más conveniente, en términos de los cálculos que se deben hacer, representar la distribución de probabilidad de una variable aleatoria de una manera algebraica. Al hacer esto, podemos llevar a cabo cálculos de probabilidad mediante la sustitución de valores numéricos directamente en una fórmula algebraica.

**Distribuciones de probabilidad**

Consideremos un espacio muestral S correspondiente a un experimento aleatorio, A cada elemento del espacio muestral, es decir, a cada resultado del experimento, podemos asociarle un número, a estos valores numéricos asociados a cada resultado forman la variable aleatoria. A cada valor de la variable aleatoria le corresponde una probabilidad. Como en el caso de las frecuencias relativas, la suma de todas las probabilidades debe ser igual a 1.

Si en un experimento aleatorio a cada suceso aleatorio elemental le asignamos un valor numérico obtenemos una variable aleatoria. Es decir, una variable que lleva asociada una probabilidad. La probabilidad de un valor concreto de la variable es la probabilidad que corresponde a los sucesos aleatorios elementales a los que hemos asignado ese valor numérico.

Resumiendo, una variable aleatoria se construye al atribuir un número (positivo, negativo o cero) a cada uno de los sucesos aleatorios que forman el espacio muestral de un experimento aleatorio. La probabilidad de cada valor de la variable es la probabilidad conjunta de los sucesos que dan lugar a ese valor. Es decir, definimos una variable aleatoria como una aplicación del espacio muestral W sobre el conjunto de los números reales R.

Según la amplitud del campo de variación de la función podemos distinguir: variables aleatorias discretas y variables aleatorias continuas. De la misma forma que en estadística descriptiva, una variable aleatoria es discreta si toma valores en un conjunto finito o infinito numerable. Y una variable aleatoria es continua si puede tomar valores en un conjunto infinito no numerable. Como ejemplo típico de variable aleatoria discreta tenemos la distribución binomial, y como ejemplo típico de variable aleatoria continua vamos a ver ahora la distribución normal.

*Por ejemplo*: supongamos que vamos a realizar un experimento aleatorio que consiste en seleccionar una persona y apuntar su peso. Podemos crear una variable aleatoria cuyos valores sean el número de kilogramos que pesa la persona observada. En este caso, el rango de valores posibles se extiende entre los límites naturales, pero la continuidad de esta variable aleatoria radica en el carácter continuo de lo que medimos, el peso, es decir, en el hecho de que entre dos valores posibles se podrían obtener infinitos valores intermedios, también posibles si utilizáramos aparatos con suficiente precisión. Estos "infinitos" en el interior del rango de la variable es lo que diferencia a las variables continuas de as discretas.

Se considera que una distribución de probabilidad es cualquier mecanismo que nos ayuda a obtener las probabilidades de los valores de una variable si es discreta, o las probabilidades de intervalos de la variable si es continua. Si la variable aleatoria es discreta es posible asignar probabilidades a cada uno de los valores puntuales de la variable. En contra, *cuando es continua cada uno de los infinitos valores posibles tendrá probabilidad cero y sólo podremos hablar de probabilidad dentro de intervalos*.

**Distribuciones de probabilidad con variable aleatoria continua, Función de Distribución y Función de Densidad.**

Si la variable aleatoria es continua hay infinitos valores posibles de la variable y entre cada dos de ellos se podrían definir infinitos valores más. En estas condiciones, y como ya hemos dicho, no es posible deducir la probabilidad de un valor puntual de la variable, como se puede hacer en el caso de variables aleatorias discretas. Pero sí es posible calcular la probabilidad acumulada hasta un cierto valor (*función de distribución*), más tarde podremos analizar cómo cambia la probabilidad acumulada en cada punto (estos cambios no son probabilidades, sino otro concepto que denominamos *densidad de probabilidad*).

*Ejemplo*

Pretendemos observar la altura de un grupo de personas y vamos a seleccionar a una persona de forma totalmente aleatoria. La probabilidad de que la altura de esa persona sea exactamente 1,62894635... m es cero. Pero la probabilidad de que la altura de esa persona esté entre 1,62 m y 1,63 m tendrá un valor concreto y casi con certeza que será mayor que la probabilidad de que esté entre 2,10 m y 2,11 m. Por tanto, la densidad de probabilidad en el entorno de 1,625 m es mayor que la densidad de probabilidad en el entorno de 2,105 m. *Sin embargo, que el valor exacto* 1,62894635 *tenga probabilidad cero de ocurrir no implica que sea imposible que ocurra.* De hecho, cualquier persona que seleccionemos tendrá una altura concreta y exacta que tenía probabilidad cero de suceder.

Al iniciar el análisis estadístico de una serie de datos, y después de la etapa de detección y corrección de errores, un primer paso consiste en describir la distribución de las variables estudiadas y, en particular, de los datos numéricos.

Una de las distribuciones teóricas mejor estudiadas en los textos de bioestadística y demás disciplinas utilizadas en la práctica, es la distribución normal, también llamada distribución gaussiana. Su importancia se debe fundamentalmente a la frecuencia con la que distintas variables asociadas a fenómenos naturales y cotidianos siguen, aproximadamente, esta distribución. Caracteres morfológicos (como la talla o el peso), o psicológicos (como el cociente intelectual) son ejemplos de variables de las que frecuentemente se asume que siguen una distribución normal.

La distribución normal fue reconocida por primera vez por el francés Abraham de Moivre (1667-1754). Posteriormente, Carl Friedrich Gauss (1777-1855) elaboró desarrollos más profundos y formuló la ecuación de la curva; de ahí que también se le conozca, más comúnmente, como la "campana de Gauss". La distribución de una variable normal está completamente determinada por dos parámetros, su media y su desviación estándar, denotadas generalmente por μ y σ.

La distribución normal es la distribución de probabilidad que con más frecuencia aparece en estadística y teoría de probabilidades. Esto se debe a dos razones fundamentalmente:

a). Su función de densidad es simétrica y con forma de campana, lo que favorece su aplicación como modelo a gran número de variables estadísticas.

b). Es además, límite de otras distribuciones y aparece relacionada con multitud de resultados ligados a la teoría de las probabilidades gracias a sus propiedades matemáticas.



La campana de Gauss o curva normal es una función de probabilidad continua, simétrica, cuyo máximo coincide pues con la media μ y la desviación típica σ