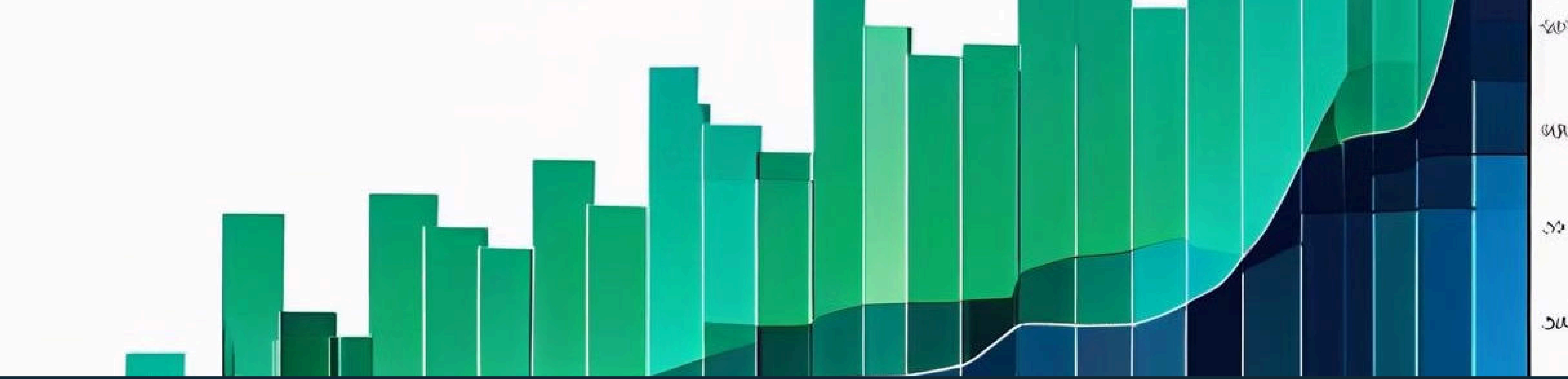
Estadísticos y Distribuciones

Entender los estadísticos clave de una muestra experimental, como la media, mediana y varianza, puede guiar la selección de la familia de distribuciones que mejor describe los datos. Sin embargo, es importante tener en cuenta que estos estadísticos calculados a partir de la muestra son observaciones aleatorias y generalmente no coincidirán con los valores verdaderos de la distribución desconocida de la cual se tomaron los datos.

Coeficiente de Variación

Un estadístico muy útil para identificar el tipo de distribución es el coeficiente de variación, definido como el cociente entre la desviación estándar y la media. Para una distribución exponencial, el coeficiente de variación siempre es 1, independientemente del valor de la media. Si el coeficiente de variación estimado de los datos experimentales está cerca de 1, es razonable suponer que la distribución es exponencial.

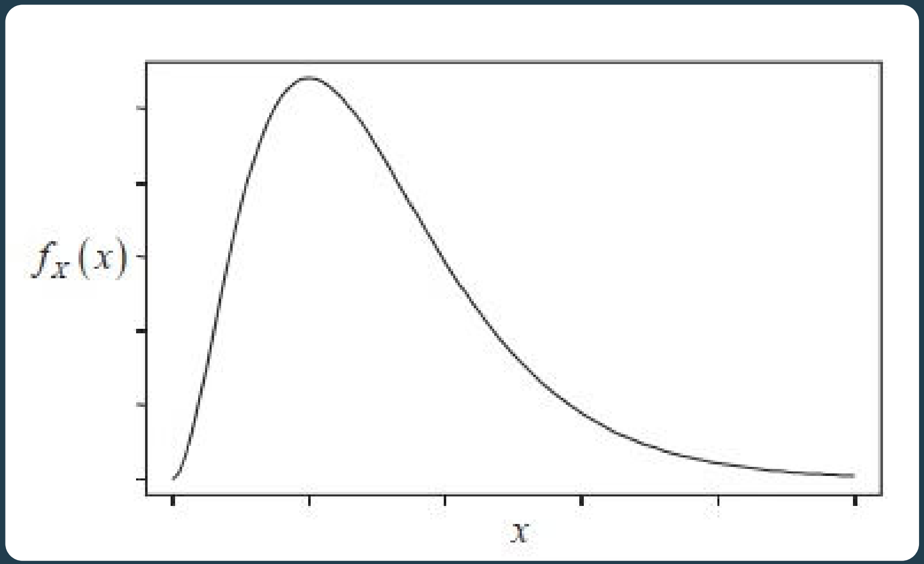


Distribuciones con Forma Asimétrica

Distribuciones Gamma y Weibull

Cuando el histograma de los datos tiene una forma similar a la mostrada, con una cola hacia la derecha, las distribuciones gamma y Weibull con parámetro de forma mayor a 1 pueden ser adecuadas. Estas distribuciones tienen un coeficiente de variación menor a 1.

Fig 4.8



Distribución Lognormal

La distribución lognormal también tiene una forma similar a la mostrada en la figura y su coeficiente de variación puede ser cualquier número real positivo. Por lo tanto, si el coeficiente de variación estimado es mayor a 1, la distribución lognormal sería más apropiada que la gamma o Weibull.

Estadísticos para Distribuciones Discretas

1

Distribución de Poisson

Para distribuciones discretas, el estadístico $\tau = \mu/\sigma^2$ juega un papel similar al coeficiente de variación en las continuas. Para la distribución de Poisson, $\tau = 1$.

2

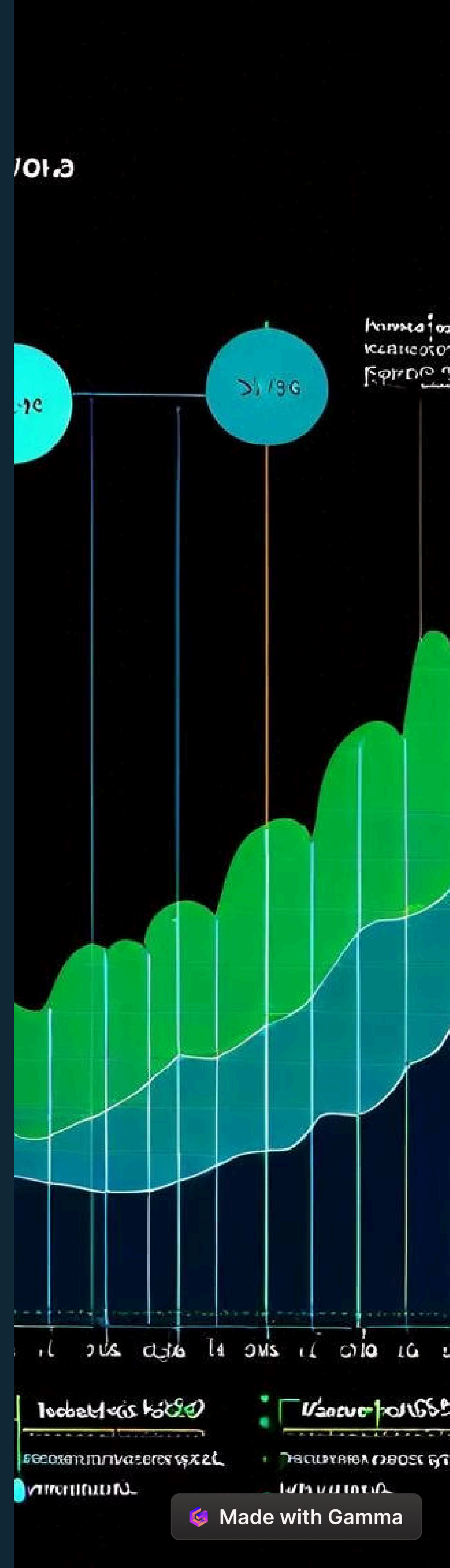
Distribución Binomial

Para la distribución binomial, $\tau < 1$, mientras que para la distribución binomial negativa, $\tau > 1$. Estos estadísticos pueden ayudar a discriminar entre estos tipos de distribuciones discretas.

3

Coeficiente de Simetría

Otro estadístico útil es el coeficiente de simetría o sesgo, que mide la simetría de la distribución. En distribuciones simétricas como la normal, el sesgo es cero.



Estimación de Estadísticos

Estimación de Coeficiente de Variación

El coeficiente de variación se puede estimar a partir de los datos experimentales como:

$$cv = \frac{\sqrt{\sigma^2}}{\mu}$$

Estimación de Tau

Para distribuciones discretas, el estadístico τ se puede estimar como:

$$\tau = \frac{\sigma^2}{\mu}$$

Estimación de Sesgo

El coeficiente de simetría o sesgo se puede estimar como:

$$\hat{\nu} = \frac{\hat{m}_3(n)}{\left(S^2(n)\right)^{3/2}}$$

$$\hat{m}_3(n) = \frac{\sum_{i:1}^{n} \left(x_i - \bar{X}(n)\right)^3}{n}$$

Ejemplo de Aplicación

1 Analizar Estadísticos

Supongamos que tenemos un conjunto de datos experimentales y queremos determinar qué tipo de distribución lo describe mejor.

2 Calcular Coeficiente de Variación

Primero, calculamos el coeficiente de variación estimado a partir de los datos. Si este valor está cerca de 1, la distribución exponencial sería una buena opción.

3 Evaluar Forma de la Distribución

Si la forma del histograma es asimétrica, con una cola hacia la derecha, las distribuciones gamma, Weibull o lognormal podrían ser adecuadas. El coeficiente de variación ayuda a distinguir entre ellas.

4 Considerar Distribuciones Discretas

Para distribuciones discretas, el estadístico τ y el coeficiente de simetría pueden ser útiles para identificar si la distribución es de Poisson, binomial o binomial negativa.

Importancia de los Estadísticos

Guía para la Selección de Distribuciones

Los estadísticos clave, como la media, mediana, varianza, coeficiente de variación y coeficiente de simetría, proporcionan información valiosa sobre el tipo de distribución que mejor se ajusta a los datos experimentales.

Estimación de Parámetros

Conocer los estadísticos de la muestra es fundamental para estimar los parámetros de la distribución subyacente, lo cual es crucial para modelar y simular procesos aleatorios de manera adecuada.

Toma de Decisiones

La elección correcta de la distribución de probabilidad, basada en el análisis de los estadísticos, permite tomar decisiones más informadas y confiables en diversos campos, como la ingeniería, la economía y la medicina.

Validación de Modelos

Comparar los estadísticos calculados a partir de los datos experimentales con los de la distribución seleccionada es una forma de validar la adecuación del modelo a la realidad observada.

Distribuciones Continuas

Distribución Normal

La distribución normal es simétrica y se caracteriza por tener media y mediana iguales.

Distribución Exponencial

La distribución exponencial tiene un coeficiente de variación igual a 1, independientemente de la media.

Distribución Gamma

La distribución gamma tiene una forma similar a la de la Figura 4.8 cuando el parámetro de forma es mayor a 1.

Distribución Weibull

La distribución Weibull también tiene una forma similar a la de la Figura 4.8 cuando el parámetro de forma es mayor a 1.

Distribuciones Discretas

1

Distribución de Poisson

La distribución de Poisson se caracteriza por tener $\tau = 1$, es decir, la media y la varianza son iguales.

2

Distribución Binomial

Para la distribución binomial, $\tau < 1$, mientras que para la distribución binomial negativa, $\tau > 1$.

3

Coeficiente de Simetría

El coeficiente de simetría o sesgo es cero en distribuciones simétricas como la normal, y diferente de cero en distribuciones asimétricas.

Conclusión

En resumen, el análisis de los estadísticos clave de una muestra experimental, como la media, mediana, varianza, coeficiente de variación y coeficiente de simetría, es fundamental para guiar la selección de la familia de distribuciones que mejor describe los datos. Estos estadísticos proporcionan información valiosa para estimar los parámetros de la distribución, validar modelos y tomar decisiones informadas en diversos campos. Comprender la relación entre los estadísticos y las diferentes distribuciones de probabilidad es un paso crucial en el modelado y simulación de procesos aleatorios.