DISTRIBUCIONES DISCRETAS DE VARIABLES ALEATORIAS

TEMA 3

Xabier Erdocia
Itsaso Leceta

BY NC SA

eman ta zabal zazu
UPV EHU

OBJETIVOS

- ✓ Ser capaz de identificar la distribución discreta que sigue la variable aleatoria
- ✓ Tras identificar la distribución discreta que sigue la variable aleatoria, ser capaz de calcular diferentes probabilidades empleando la función de probabilidad o distribución
- ✓ Poseer la capacidad de calcular e interpretar los momentos de diferentes distribuciones discretas
- ✓ Conocer las condiciones que son necesarias para aproximar la distribución hipergeométrica mediante la distribución binomial y la distribución binomial mediante la distribución de Poisson y ser capaz de realizar la aproximación correctamente

ÍNDICE

OCW

eman ta zabal zazu

UPV EHU

3.1. Distribución binaria

- Cuando solamente existen dos posibles resultados al realizar un único ensayo aleatorio, la distribución que sigue la variable aleatoria es una distribución binaria o de Bernoulli.

- Los dos posibles resultados son $X=1$ (Éxito) o $X=0$ (Fracaso), siendo las probabilidades $P(X=1)=p$ y $P(X=0)=q=1-p$.

- La distribución que sigue la variable se expresa como:

$$X \sim \text{Binaria}(p)$$

- La función de probabilidad es la siguiente:

$$p(x)=p^x(1-p)^{1-x}=p^x q^{1-x}$$

- La función de distribución es la siguiente:

$$F(x)=P(X\le x)=\begin{cases} q & x=0 \\ 1 & x=1 \end{cases}$$

- La media de una variable aleatoria que sigue una distribución binaria es:

$$E(X) = p$$

- Conociendo la función de probabilidad, se pueden calcular otros momentos y medidas. Por ejemplo, el valor de la varianza es:

$$\sigma^2 = p \cdot q$$

3.2. Distribución binomial

eman ta zabal zazu

UPV EHU

- Si se repite en n ocasiones el ensayo aleatorio realizado en la distribución binaria (siendo cada uno de los ensayos independiente del resto), la distribución que sigue la variable aleatoria que considera los resultados obtenidos en todos lo ensayos es una distribución binomial.

- En cada ensayo hay dos posibles resultados, “Éxito” con una probabilidad p y “Fracaso” con una probabilidad $q = 1 - p$.

- La distribución que sigue la variable se expresa como:

$$X \sim \mathrm{B}(n, p)$$

- La función de probabilidad es la siguiente:

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n$$

- La función de distribución es la siguiente:

$$F(x) = P(X \leq x) = \sum_{i=0}^{x} p(i) = \sum_{i=0}^{x} \binom{n}{i} p^i q^{n-i}$$

- La media de una variable aleatoria que sigue una distribución binomial es:

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

- Conociendo la función de probabilidad, se pueden calcular otros momentos y medidas. Por ejemplo, el valor de la varianza es:

$$\sigma^2 = n \cdot p \cdot q$$

3.3. Distribución geométrica

- Teniendo en cuenta el ensayo aleatorio definido en la distribución binaria (siendo cada ensayo independiente del resto), la distribución que sigue la variable aleatoria que considera el número de ensayos realizados hasta obtener el primer éxito (sin tener en cuenta éste) es una distribución geométrica.

- En cada ensayo hay dos posibles resultados, "Éxito" con una probabilidad p y "Fracaso" con una probabilidad $q = 1 - p$.

- La distribución que sigue la variable se expresa como:

$$X \sim \mathrm{G}(p)$$

- La función de probabilidad es la siguiente:

$$p(x) = q^x p \qquad x = 0, 1, 2, \ldots$$

- La función de distribución es la siguiente:

$$F(x) = P(X \leq x) = \sum_{i=0}^{x} p(i) = \sum_{i=0}^{x} q^i p$$

- La media de una variable aleatoria que sigue una distribución geométrica es:

$$E(X) = \frac{q}{p}$$

- Conociendo la función de probabilidad, se pueden calcular otros momentos y medidas. Por ejemplo, el valor de la varianza es:

$$\sigma^2 = \frac{q}{p^2}$$

3.4. Distribución binomial negativa

- Teniendo en cuenta el ensayo aleatorio definido en la distribución binaria (siendo cada ensayo independiente del resto), la distribución que sigue la variable aleatoria que considera el número de ensayos realizados hasta obtener *n* éxitos (sin tener en cuenta los *n* éxitos) es una distribución binomial negativa.

- En cada ensayo hay dos posibles resultados, "Éxito" con una probabilidad p y "Fracaso" con una probabilidad $q = 1 - p$.

- La distribución que sigue la variable se expresa como:

$$X \sim \mathrm{BN}(n, p)$$

- La función de probabilidad es la siguiente:

$$p(x) = \binom{n+x-1}{x} q^x p^n \qquad x = 0, 1, 2, \ldots$$

- La función de distribución es la siguiente:

$$F(x) = P(X \leq x) = \sum_{i=0}^{x} p(i) = \sum_{i=0}^{x} \binom{n+i-1}{i} q^i p^n$$

- La media de una variable aleatoria que sigue una distribución binomial negativa es:

$$E(X) = \frac{nq}{p}$$

- Conociendo la función de probabilidad, se pueden calcular otros momentos y medidas. Por ejemplo, el valor de la varianza es:

$$\sigma^2 = \frac{nq}{p^2}$$

3.5. Distribución hipergeométrica

- La distribución hipergeométrica es similar a la distribución binomial pero el muestreo no es independiente, esto es, las observaciones se realizan en una población finita sin reemplazamiento.

- En una población finita, no devolver un elemento a la población tras observarlo, influye en las siguientes observaciones. Las proporciones no se mantienen constantes a medida que las observaciones avanzan.

- En cada observación de una población de tamaño N hay dos posibles resultados, "Éxito" y "Fracaso". En la primera observación (ensayo) la probabilidad de éxito es p y la probabilidad de fracaso es $q = 1 - p$.

- La distribución que sigue la variable se expresa como:

$$X \sim \mathrm{H}(N, n, p)$$

- La función de probabilidad es la siguiente:

$$p(x) = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} \qquad \max(0, n - Nq) \le x \le \min(n, Np)$$

- La función de distribución es la siguiente:

$$F(x) = P(X \leq x) = \sum_{i=\max(0,n-Nq)}^{x} p(i) = \sum_{i=\max(0,n-Nq)}^{x} \frac{\binom{Np}{i}\binom{Nq}{n-i}}{\binom{N}{n}}$$

- La media de una variable aleatoria que sigue una distribución hipergeométrica es:

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

- Conociendo la función de probabilidad, se pueden calcular otros momentos y medidas. Por ejemplo, el valor de la varianza es:

$$\sigma^2 = n \cdot p \cdot q \frac{(N-n)}{(N-1)}$$

Aproximación de una distribución hipergeométrica mediante una distribución binomial:

- En una variable aleatoria que sigue una distribución hipergeométrica, si el tamaño de la población, *N*, es mucho mayor que el número de observaciones realizadas, *n*, esta distribución hipergeométrica puede aproximarse mediante una distribución binomial.

- Por consiguiente, si $N \to \infty$ y $n \to 0$, $H(N, n, p) \cong B(n, p)$

- Esta aproximación es aceptable cuando $N > 10 \cdot n$

3.6. Distribución de Poisson

- Cuando una variable aleatoria X define el número de veces que un evento se repite en un intervalo continuo (tiempo, espacio, ...), ésta sigue una distribución de Poisson. Los eventos que ocurren en un intervalo son independientes de aquellos que se producen en otro intervalo, siempre y cuando los intervalos no se superpongan.

- El número de eventos es siempre positivo y el parámetro λ definirá el número de eventos esperados en un intervalo dado.

- La distribución que sigue la variable se expresa como:

$$X \sim \mathscr{P}(\lambda)$$

- La función de probabilidad es la siguiente:

$$p(x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!} \qquad \lambda, x \geq 0$$

- La función de distribución es la siguiente:

$$F(x) = P(X \leq x) = \sum_{i=0}^{x} p(i) = \sum_{i=0}^{x} \frac{e^{-\lambda} \cdot \lambda^{i}}{i!}$$

OCW

eman ta zabal zazu

UPV EHU

- La media de una variable aleatoria que sigue una distribución de Poisson es:

$$E(X) = \lambda$$

- Conociendo la función de probabilidad, se pueden calcular otros momentos y medidas. Por ejemplo, el valor de la varianza es:

$$\sigma^2 = \lambda$$

Aproximación de una distribución binomial mediante una distribución de Poisson:

- En una variable aleatoria que sigue una distribución binomial, cuando el número de observaciones o ensayos, *n*, es muy elevado y la probabilidad de éxito, *p*, es baja, la distribución binomial puede aproximarse mediante una distribución de Poisson.

- Por consiguiente, si $n \to \infty$ y $p \to 0$, $B(n, p) \cong \mathscr{P}(n \cdot p = \lambda)$

- Esta aproximación es aceptable cuando $n > 30$ y $p \leq 0{,}1$

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea