Universidad de Santander UDES

MAGNA EDUCATIO AD MELLIOR

UNIVERSIDAD DE SANTANDER

VIGILADA MINEDUCACIÓN | SNIES 2832

Estadística

Variable aleatoria discreta

Daniel Martínez Bello

Universidad de Santander
Maestría Biotecnología

Octubre de 2025

Variable aleatoria

- Una **variable aleatoria** es el resultado numérico de un experimento aleatorio.
- Una variable aleatoria corresponde a todos los posibles resultados del experimento junto con la probabilidad de cada resultado.
- Se encuentran variables aleatorias **discretas** o **continuas**.
- La variable aleatoria discreta corresponde a una variable aleatoria que solo considera un número contable de posibilidades.
  - $P(X = k)$
- La variable aleatoria continua puede tomar cualquier valor en la linea real o en un subconjunto de la línea real.
  - $P(X \in A)$

Variable aleatoria

- Dado un experimento aleatorio, el espacio muestral es el conjunto de todos los resultados posibles del experimento aleatorio.
- El espacio muestral se denota como Ω.
- Los elementos de Ω son los puntos muestrales ω.
- Un evento es todo subconjunto E de Ω, $E \in \Omega$.
- El conjunto de todos los eventos es el conjunto de partes de Ω, $\mathcal{P}(\Omega)$.
- Un evento cualquiera E ocurre cuando ocurre ω, esto es, $\omega \in E$.
- Entre todos los eventos, existe el evento imposible $\emptyset$, y el evento seguro Ω.

Variable aleatoria

- Un subconjunto α de $\mathcal{P}(\Omega)$, no vacio, que verifique, (1) el complementario de un evento de α tambien pertenece a α; y (2) la unión de una colección infinita numerable de eventos de α también pertenece a α, se denomina una **$\boldsymbol{\sigma}$-algebra** de Boole.
- El par (Ω, α) se denomina **espacio probabilizable**.
- La probabilidad es una medida relativa y teórica de la ocurrencia de un evento.
- Dado un espacio probabilizable (Ω, α), se denomina **probabilidad** sobre α a una aplicación $p : \alpha \to \mathbb{R}$ que verifica
  - $p(E) \geq 0$, para todo $E \in \alpha$.
  - $p(\Omega) = 1$.
  - $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$, para toda colección infinita numerable de eventos E_i de α, disjuntos de dos a dos.

Variable aleatoria

- La terna (Ω, α, p) se denomina **espacio probabilístico**.
- Dado un (Ω, α, p), una variable aleatoria es una aplicación $X : \Omega \to \mathbb{R}$ si

$$X^{-1}((-\infty, x]) \in \alpha \text{ para todo número real } x \tag{1}$$

- El conjunto $X^{-1}((-\infty, x])$ lo podemos denotar $[X \leq x]$.
- Una variable aleatoria secaracteriza no solo por intervalos de la forma $(-\infty, x])$ sino también por cualquier intervalo.
- Una variable aleatoria es una aplicación de $X : \Omega \to \mathbb{R}$ que verifica que $X^{-1}(I) \in \alpha$, para todo intervalo real I.

Ejemplos de variables aleatorias

- El resultado de un lanzamiento al aire de una moneda.
- El resultado de lanzar un dado
- Extraer una carta al azar de una baraja
- La glucosa en sangre de una persona.
- El tiempo al momento en que una persona muere de cancer despues de ser diagnosticada
- La decisión de adquirir un producto por parte de un consumidor

Función de masa de probabilidad

La función de masa de probabilidad (FMP) evaluada en un valor corresponde a la probabilidad que la variable aleatoria tome ese valor.

Una FMP p debe satisfacer,

1. $p(x) \geq 0$ para todo x
2. $\sum_x p(x) = 1$

La suma se toma sobre todos los posibles valores de x.

Ejemplo

Sea X el resultado de un lanzamiento de una moneda donde $X = 0$ representa sello y $X = 1$ representa cara.

$$p(x) = (1/2)^x (1/2)^{1-x} \quad \text{para} \quad x = 0, 1$$

Supongamos que no sabemos si la moneda es justa o no; Sea θ la probabilidad de cara expresada como una proporción, esto es (entre 0 y 1).

$$p(x) = \theta^x (1-\theta)^{1-x} \quad \text{para} \quad x = 0, 1$$

Ejemplo

Para la moneda con trampa

$$p(0) = 1 - \theta \text{ y } p(1) = \theta$$

asi

$$p(x) > 0 \text{ para } x = 0, 1$$

y

$$p(0) + p(1) = \theta + (1 - \theta) = 1$$

Función de distribución acumulada y función de sobrevida

- La **función de distribución acumulada** (FDA) de una variables aleatoria X se define como la función
$$F(x) = P(X \leq x)$$
- Esta definición se aplica independiente de si X es discreta o continua
- La **función de sobrevida** de una variable aleatoria X se define como
$$S(x) = P(X > x)$$
$$S(x) = 1 - F(x)$$
- Para las variables aleatorias continuas, la FDP es la derivada de la FDA
$$f(x) = \frac{d}{dx}F(x)$$

Propiedades de la FDA

- $0 \leq F(x) \leq 1$, para todo número real x.
- F es monótona creciente.
- $\text{Lím}_{x\to\infty} F(x) = 1$ y $\text{Lím}_{x\to-\infty} F(x) = 0$
- $P(a < X \leq b) = F(b) - F(a)$
- F es continua por la derecha.

Cuantiles

- El α^{esimo} **cuantil** de una distribución con función de distribución F es el punto x_α de tal forma que

$$F(x_\alpha) = \alpha$$

- Un **percentil** es un cuantil con α expresado como porcentaje
- La **mediana** es el 50^{esimo} percentil

Modelos de probabilidad

- Los modelos de probabilidad conectan los datos a la población a traves de unos supuestos.
- El investigador asume un modelo de probabilidad, y genera inferencia, es decir, a partir de muestras al azar, se obtiene información de la población.

Valor esperado

- El **valor esperado** o esperanza matemática de una variable aleatoria es el centro de su distribución
- Para una variable aleatoria discreta X con FMP $p(x)$, se define como

$$E[X] = \sum_{x} xp(x).$$

Donde la suma se toma sobre todos los valores posibles x

- $E[X]$ representa el centro de gravedad de un conjunto de eventos con su ponderación, $\{x, p(x)\}$

Ejemplo

- X es una variable aleatoria que representa los resultados del experimento aleatorio lanzar un dado y ver el número que muestra el dado
- Cuál es el valor esperado de X?

$$E[X] = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3.5$$

Reglas sobre los valores esperados

- El valor esperado es un operador lineal
- Si a y b no son aleatorios X y Y son dos variables aleatorias, entonces
  - $E[aX + b] = aE[X] + b$
  - $E[X + Y] = E[X] + E[Y]$
- *En general* si g es una función que no es lineal

$$E[g(X)] \neq g(E[X])$$

- Por ejemplo, $E[X^2] \neq E[X]^2$

Ejemplo

- Se lanza un dado dos veces. Cuál es el valor esperado del promedio?
- Dado X_1 y X_2 los resultados de los dos lanzamientos

$$E[(X_1 + X_2)/2] = \frac{1}{2}(E[X_1] + E[X_2]) = \frac{1}{2}(3.5 + 3.5) = 3.5$$

La varianza

- La varianza de una variable aleatoria es una medida de *dispersión*
- Si X es una variable aleatoria con media μ, la varianza de X se define como

$$\mathrm{Var}(X) = E[(X-\mu)^2]$$

  Que corresponde al cuadrado de la distancia de X a la media
- Densidades con varianzas grandes estan mas dispersas que densidades con un varianza baja
- Otra forma de expresarlo

$$\mathrm{Var}(X) = E[X^2] - E[X]^2$$

- Si a es una constante, $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$
- La raíz cuadrada de la varianza se llama la **desviación estándar**
- La desviación estándar tiene las mismas unidades que X

Propiedades de la varianza de la variable aleatoria

- $\text{Var}(X)$ es un número.
- $\text{Var}(c) = 0$. Varianza de una constante c es igual a 0.
- $\text{Var}(X + c) = \text{Var}(X)$. Varianza de la suma de una constante con una variable aleatoria es igual a la varianza de la variable aleatoria.
- $\text{Var}(cX) = c^2\text{Var}(X)$. Varianza del producto de una constante c por una variable aleatoria es igual a la varianza de la variable aleatoria multiplicada por el cuadrado de la constante.

- **Ejemplo**: Cuál es la varianza de la muestra del resultado del lanzamiento de un dado?
  - $E[X] = 3.5$
  - $E[X^2] = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = 15.17$
- $\mathrm{Var}(X) = E[X^2] - E[X]^2 \approx 2.92$
- **Ejemplo**: Cuál es la varianza de la muestra del resultado de lanzar una moneda al aire con probabilidad de cara p?
  - $E[X] = 0 \times (1-p) + 1 \times p = p$
  - $E[X^2] = E[X] = p$
- $\mathrm{Var}(X) = E[X^2] - E[X]^2 = p - p^2 = p(1-p)$
- **Ejemplo**: Suponiendo que una variable aleatoria es tal que $0 \leq X \leq 1$ y $E[X] = p$
- $X^2 \leq X$, así $E[X^2] \leq E[X] = p$
- $\mathrm{Var}(X) = E[X^2] - E[X]^2 \leq E[X] - E[X]^2 = p(1-p)$
- Debido a esto, la varianza de Bernoulli es la mas larga posible para las variables aleatorias acotadas entre 0 y 1

Distribución degenerada

Como un caso extremo, se puede considerar un experimento aleatorio que genera solo un posible resultado x_0 con probabilidad 1.

$$
\begin{aligned}
X &\sim \text{ Degenerada en } x_0 \\
P(X = x) &= 1 \text{ Para } x = x_0 \\
E(X) &= x_0 \\
Var(X) &= 0
\end{aligned}
$$

Distribución Uniforme

Una distribución uniforme discreta corresponde a un mecanismo que genera cada uno de sus N diferentes resultados con la misma probabilidad. Los N diferentes resultados se denotan por $1, \cdots, N$. Un ejemplo típico corresponde a los resultados y la probabilidad de cada resultado del lanzamiento de un dado.

$$X \sim \text{Uniforme sobre } 1, 2, \cdots, N$$

$$P(X = x) = \frac{1}{N} \text{ para } x = 1, 2, \cdots, N$$

$$E(X) = \frac{N+1}{2}$$

$$\text{Var}(X) = \frac{N^2 - 1}{12}$$

Distribución Bernoulli

Suponga que un experimento aleatorio genera solo dos tipos de resultado, tales como si/no, o blanco/negro, o gana/pierde, o exito/fracaso. Recodificando los posibles resultados a 1 ó 0, y denotando la probabilidad de generar un uno por π. El ejemplo típico es lanzar una moneda al aire.

$$X \sim \text{Bernoulli}(1;\pi) \quad (0 < \pi < 1)$$

$$P(X = x) = \begin{cases} \pi & \text{para } x = 1 \\ (1-\pi) & \text{para } x = 0 \end{cases}$$

$$E(X) = \pi$$

$$Var(X) = \pi(1-\pi)$$

Distribución Binomial

Un experimento aleatorio típico es como sigue. Defina n de antemano. De una urna marcada con fichas 0-1 que contiene N fichas, de las cuales s son unos,y $N - s$ son ceros. Denote $\frac{s}{N}$ por π. Obtenga una muestra al azar y con reemplazamiento de n fichas de esta urna, de tal forma que las muestras siguientes sean independientes. Llame a X el número de unos observado en las n fichas muestreadas.

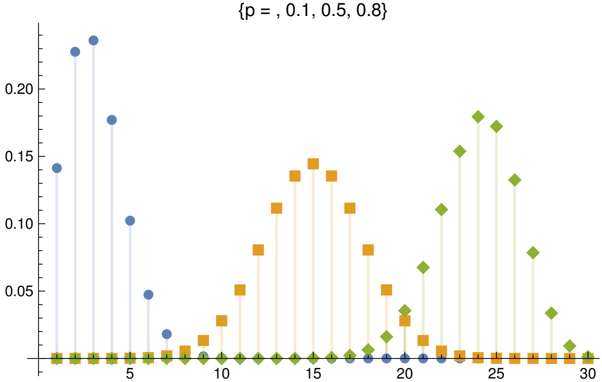
$$X \sim Bin(n, \pi)$$

$$P(X = x) = \binom{n}{x} \pi^x (1-\pi)^{n-x} \quad \text{para } x = 0, 1, \cdots, n$$

$$\begin{aligned} E(X) &= n\pi \\ Var(X) &= n\pi(1-\pi) \\ Moda(X) &= (n+1)\pi \end{aligned}$$

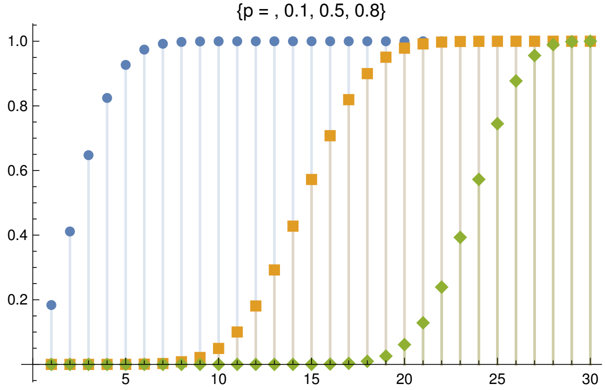
Distribución Binomial

Función de densidad



Distribución Binomial

Función de distribución acumulada



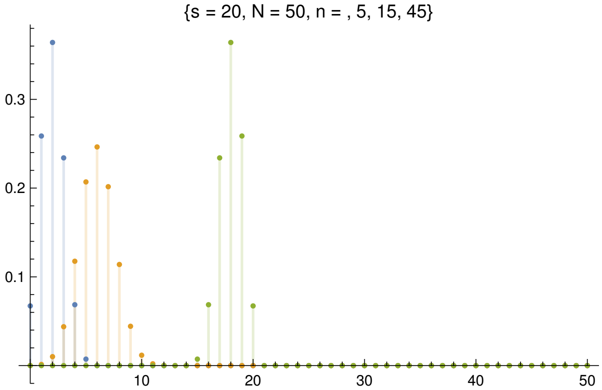
Distribución hipergeométrica

Un experimento aleatorio es como sigue. Defina n de antemano. De una urna marcada con fichas 0-1 que contiene N fichas, de las cuales s son unos,y $N - s$ son ceros. Denote $\frac{s}{N}$ por π. Seleccione una muestra al azar y sin reemplazamiento de n fichas de esta urna. Defina como X el número total de unos observado en las n fichas.

$$
\begin{aligned}
X &\sim \text{Hipergeométrica}(N; n; \pi) \\
P(X = x) &= \frac{\binom{s}{x}\binom{N-s}{n-x}}{\binom{N}{n}} \text{para } \max(0, s + n - N) \leq x \leq \min(s, n) \\
E(X) &= n\pi \\
\text{Var}(X) &= \frac{N-n}{N-1} \cdot \pi(1-\pi)
\end{aligned}
$$

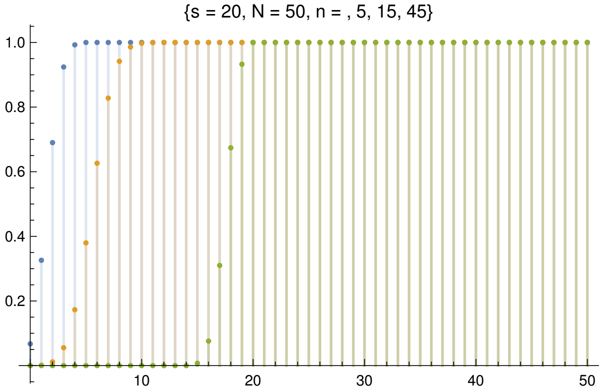
Distribución Hipergeométrica

Función de densidad



Distribución Hipergeométrica

Función de distribución acumulada



Distribución hipergeométrica

- Un fabricante de empanadas informa que en un envio de 150 empanadas, 80 son de pollo. Si se toma una muestra al azar de 20 empanadas, cuál es la probabilidad de obtener exactamente 4 empanadas de pollo
- Si X es el número de empanadas de pollo en las 20 empanadas, X es hipergeométrica con N=150, n=20,k=80

$$P(X=4)=\frac{\binom{80}{4}\binom{70}{16}}{\binom{150}{20}}=0.00108$$

Distribución hipergeométrica

Una fundación tiene 60 miembros; 25 son hombres y 35 son mujeres. Se elige al azar un comité de 8 miembros. Encuentra cada uno de los siguientes:

- La función de densidad de probabilidad del número de mujeres en el comité.
- La probabilidad de que los miembros del comité sean todos del mismo género

Un pequeño estanque contiene 2000 peces; 150 están etiquetados. Supongamos que se capturan 20 peces. Encuentra cada uno de los siguientes:

- La función de densidad de probabilidad del número de peces marcados en la muestra.
- La probabilidad de que la muestra contenga al menos 2 peces etiquetados.

Distribución Poisson

La distribución de Poisson es un mecanismo de azar que es importante para modelar la ocurrencia de eventos aleatorios en tiempo y espacio. Los supuestos de este modelo de probabilidad son: que las condiciones permanecen constantes en el tiempo, y que los intervalos de tiempo no se solapan, en el sentido que la información de los eventos en un intervalo de tiempo no revela información sobre los eventos en otro intervalo de tiempo. Un experimento aleatorio de Poisson tiene la característica denominada intensidad con la cuál el evento ocurre por unidad de tiempo, y se denota como λ

$$X \sim \text{Poisson}(\lambda)$$

$$P(X = x) = \frac{1}{x!}\lambda^x e^{-\lambda} \text{ para } x = 0, 1, 2, \ldots$$

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

$$Moda(X) = \lambda$$

Distribución Poisson

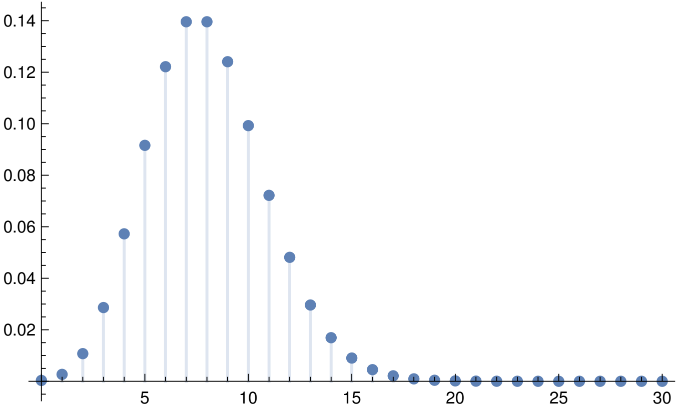
- Una persona trabaja en un call-center, y de experiencias pasadas se sabe que típicamente se tienen 8 llamadas por hora en promedio.
- Suponiendo que no se pueden tener dos llamadas al mismo tiempo, el modelo de probabilidad para X que es la variable aleatoria que corresponde al número de llamadas por hora, puede tener una distribución de Poisson con parámetro λ igual a 8 por hora.

$$P(X = x) = \frac{e^{-8}8^x}{x!}, \text{ para } x = 0, 1, 2, 3, \ldots$$

- La probabilidad de obtener 20 llamadas es igual a

$$P(X = 20) = \frac{e^{-8}8^{20}}{20!} = 0.000093$$

Poisson con $\lambda = 8$



Distribución Geométrica

Se realiza un experimento Bernoulli (probabilidad de exito $= \pi$) varias veces y bajo las mismas condiciones, hasta observar el primer exito. Defina X como el número de fallos consecutivos hasta el primer exito.

$$X \sim \text{Geométrica}(\pi) \quad (0 < \pi < 1)$$

$$P(X = x) = (1-\pi)^x \pi \text{ para } x = 0, 1, 2, \ldots$$

$$E(X) = \frac{1-\pi}{\pi}$$

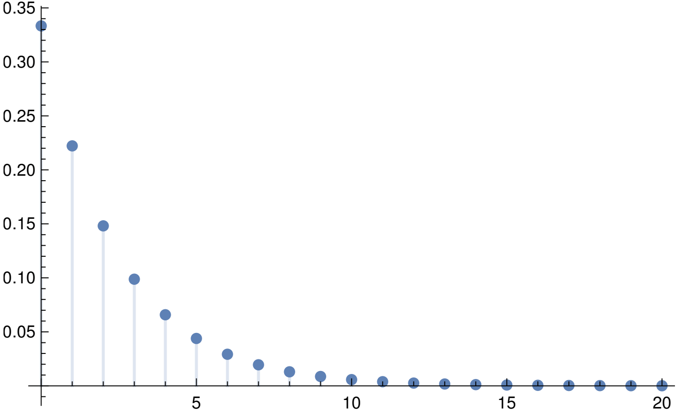
$$\text{Var}(X) = \frac{1-\pi}{\pi^2}$$

Distribución Geométrica

- Usted esta esperando en una esquina a que pase un taxi, y se pone a contar el número de automoviles que pasan hasta que pasa el primer taxi.
- Se fija el número de éxitos (que pase un taxi) pero no el número de de conteos.
- Supongamos que en el barrio se tiene probabilidad de 1/3 de ver 6 autos antes de que pase un taxi por la esquina de su casa

$$P(X = x) = (1 - \frac{1}{3})^6 \frac{1}{3} = 0.0292$$

Distribución Geométrica



Parámetro igual a $1/3$

Distribución Binomial negativa

Se realiza un experimento Bernoulli (probabilidad de exito $= \pi$) varias veces y bajo las mismas condiciones, es decir muestreo con reemplazamiento. Se define la variable aleatoria X que denota el número de fallos x antes del n-ésimo éxito. También se caracteriza como el tiempo de espera hasta el n-ésimo éxito. Donde x es el número de fallos y n es el número de exitos y π es la probabilidad de exito.

$$\begin{aligned} X \sim & \text{Neg-bin}(x; n, \pi) \\ P(X = x) = & \binom{x+n-1}{x} \pi^n (1-\pi)^x \\ E(X) = & \frac{n(1-\pi)}{\pi} \\ \text{Var}(X) = & \frac{n(1-\pi)}{\pi^2} \end{aligned}$$

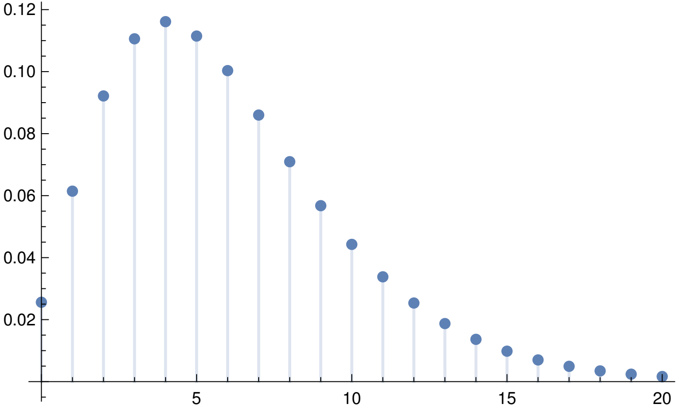
Distribución Binomial negativa

- Un jugador de basketball tiene un promedio de encestar de 0.40.
- El jugador quiere lanzar el balón hasta lograr cuatro cestas en total (no se requiere que las cuatro cestas sean seguidas una detras de la otra).
- Calcular la probabilidad de fallar seis veces antes de completar las cuatro cestas.
- Número de fallas $= 6 = x$, número de exitos esperado (n) es igual a 4, y la probabilidad (π) es 0.4

$$P(X = 6) = \binom{6 + 4 - 1}{6} 0.4^4 (1 - 0.4)^6 = 0.024$$

Distribución Binomial negativa

Binomial negativa con parámetro igual a π=0.4 y x =6



- Una urna contiene 3 pelotas rojas 6 verdes. Calcule la probabilidad de extraer 5 pelotas rojas, después de extraer 7 pelotas verdes.
- Suponga que la probabilidad de un nacimiento de un niño varón es 0.5. Una pareja desea tener exactamente dos niñas en su familia. Tendrán hijos hasta que esta condición se satisfaga.
  - ¿Cuál es la probabilidad de que la familia tenga x varones?
  - ¿Cuál es la probabilidad de que la familia tenga cuatro hijos?
  - ¿Cuál es la probabilidad de que la familia tenga cuando mucho cuatro hijos?
  - ¿Cuántos varones cree que tendría esta familia? ¿Cuántos hijos esperaría que tenga esta familia?