Aula360
Aprendizaje para el siglo XXI

Apuntes Unidad 3

Coeficiente binomial y distribución binomial

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS UNIVERSIDAD DE CHILE

CMM Centro de Modelamiento Matemático

En esta lección estudiaremos acerca de una distribución en particular y esta es la **Distribución Bernoulli.**

DISTRIBUCIÓN BERNOULLI

Una variable aleatoria de Bernoulli, llamada así por el matemático suizo Jacob Bernoulli, se asocia a un experimento aleatorio que sólo puede obtener dos resultados, usualmente nombrados como **éxito** y **fracaso**.

Si definimos una variable aleatoria discreta X que asocie el resultado "éxito" al valor 1 y el resultado "fracaso" al valor 0, se dice que la variable X tiene una distribución de Bernoulli, para la cual el valor 1 ocurre con probabilidad p, mientras que el valor 0 sucede con probabilidad $(1-p)$.

En términos generales, se dice que la variable aleatoria X se distribuye como una Bernoulli de parámetro p, con $0 \leq p \leq 1$ y se escribe de la siguiente manera:

$$X \sim Bernoulli(p)$$

A partir de esta definición, podemos expresar la función de probabilidad de una variable aleatoria discreta Xque se distribuye como una Bernoulli como:

$$P(\text{Éxito}) = P(X=1) = p$$
$$P(\text{Fracaso}) = P(X=0) = (1-p)$$

Observación

Hay muchos experimentos en los que hay más de dos posibles resultados, por lo que no es natural considerar una variable aleatoria de tipo Bernoulli. Sin embargo, siempre es posible definir una considerando algún resultado o conjunto de ellos como "éxito" y los restantes resultados como "fracaso".

Por ejemplo, al lanzar un dado podemos considerar como éxito obtener un 4 y fracaso que salga cualquier otro valor, o bien que el éxito y fracaso sea obtener un par o impar, respectivamente.

Ahora que hemos revisado esta distribución, estudiemos su valor esperado y su varianza.

VALOR ESPERADO Y VARIANZA DE UNA VARIABLE ALEATORIA DE BERNOULLI

El valor esperado y la varianza de una variable aleatoria X de tipo Bernoulli tienen un valor preciso que depende del parámetro p, correspondiente a la probabilidad de éxito. Si calculamos el valor esperado de esta variable, obtenemos lo siguiente:

$$\begin{aligned} \mathrm{E}(X) &= 0 \cdot \mathrm{P}(X=0) + 1 \cdot \mathrm{P}(X=1) \\ &= 0 \cdot (1-p) + 1 \cdot p \\ &= p \end{aligned}$$

Es decir, el valor esperado es igual a la probabilidad de éxito. Por otro lado, para calcular la varianza, podemos usar la expresión:

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

Dado que ya sabemos cuánto vale $E(X)$, basta calcular el valor de $E(X^2)$:

$$\begin{aligned} \mathrm{E}(X^2) &= 0^2 \cdot \mathrm{P}(X=0) + 1^2 \cdot \mathrm{P}(X=1) \\ &= 0 \cdot (1-p) + 1 \cdot p \\ &= p \end{aligned}$$

Por lo tanto, la varianza de la variable X está dada por:

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{E}(X^2) - (\mathrm{E}(X))^2 \\ &= p - p^2 \\ &= p \cdot (1-p) \end{aligned}$$

Antes de estudiar los conceptos vistos en esta lección, debemos tener en cuenta algunas definiciones previas. Recordemos que un experimento aleatorio es de tipo Bernoulli si sólo puede obtener dos resultados, los cuales se denominan como "éxito" y "fracaso". La variable aleatoria asociada a este experimento toma el valor 1 con probabilidad p y el valor 0 con probabilidad $(1 - p)$. Además, debemos tener presente que para representar los resultados de un experimento aleatorio, puede resultar útil trabajar con $k - tuplas$, que son listas ordenadas de k elementos. Por ejemplo, (S, C, C) es una 3-tupla que representa los resultados de lanzar 3 veces una moneda. Ahora, para contar la cantidad de k -tuplas que se pueden formar con los elementos de un conjunto de tamaño n, hay que distinguir si los elementos en ellas pueden o no repetirse:

- Cuando la $k - tupla$ puede tener elementos repetidos, el número total de $k - tuplas$ distintas es n^k. A estas $k - tuplas$ se les conoce como **variaciones con reposición.**

- Cuando la $k - tupla$ no puede tener elementos repetidos, el número total de $k - tuplas$ distintas es $n \cdot (n - 1) \cdot (n - 2) \cdot ... \cdot (n - k + 1)$. A estas $k - tuplas$ se les conoce como **variaciones sin reposición.**

Por último, recordemos que el factorial de un número n se denota por n! y corresponde al producto de los primeros n naturales, es decir, $n! = n \cdot (n - 1) \cdot ... \cdot 2 \cdot 1$. Además, por convención se define $0! = 1$.

En esta lección introduciremos un par de conceptos que nos permitirán trabajar de mejor manera con los conjuntos.

NÚMERO COMBINATORIO

Al tratar de resolver problemas de probabilidades es usual que surja la necesidad de saber cuántos subconjuntos de cierto tamaño pueden formarse con los elementos de un conjunto. A estos subconjuntos se les llama **combinaciones**. En general, las combinaciones suelen aparecer al querer contar los posibles resultados de experimentos aleatorios donde los objetos son distinguibles entre sí y estos se extraen sin reposición.

A la cantidad de combinaciones distintas de tamaño k cuyos elementos se eligen de un conjunto con n elementos lo llamaremos **número combinatorio** y lo denotaremos como:

Curso: Probabilidad y estadística descriptiva
Unidad 3 : Situaciones o fenómenos que se modelan por medio de las distribuciones binomial y normal
Tema: Distribución binomial
Contenido: Coeficiente binomial y distribución binomial

$$\binom{n}{k}$$

Cuando $k = n$, estas tuplas se conocen como **permutaciones**, y es frecuente que aparezcan en los problemas de conteo, ya que dan cuenta de todas las posibles formas en que pueden ordenarse n elementos de un conjunto en una $k - tupla$. Notemos que la fórmula anterior, nos dice que el número de permutaciones de n elementos es

$$n! = n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdot \ldots \cdot 2 \cdot 1$$

Una estrategia de conteo para calcular el número de combinaciones de tamaño k de un conjunto con n elementos es a partir de las variaciones sin reposición que hay del mismo tamaño. Si listamos las $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)$ variaciones, o formas de

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-k+1)}{k!}$$

elegir ordenadamente k elementos de un conjunto con n elementos, cada combinación aparece representada $k!$ veces en la lista.
Por otro lado, se puede calcular utilizando la siguiente expresión:
Además, es importante mencionar que el número combinatorio cumple las siguientes propiedades:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

- **Propiedad 1.** Para todo n, se cumple que

$$\binom{n}{0} = 1 \qquad \binom{n}{n} = 1$$

- **Propiedad 2.** Para todo n y k, con $n \geq 1$ y $1 \leq k \leq n-1$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Curso: Probabilidad y estadística descriptiva
Unidad 3 : Situaciones o fenómenos que se modelan por medio de las distribuciones binomial y normal
Tema: Distribución binomial
Contenido: Coeficiente binomial y distribución binomial

TRIÁNGULO DE PASCAL

El triángulo de Pascal es una representación en la que los números combinatorios se disponen en forma de triángulo, de tal forma que la fila n −ésima corresponde a los valores

$$
\begin{array}{ccccccccccc}
 & & & & & \binom{0}{0} & & & & & \\
 & & & & \binom{1}{0} & & \binom{1}{1} & & & & \\
 & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & \\
 & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & \\
 & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & \\
\binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5}
\end{array}
\qquad
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

de los números $\binom{n}{k}$, con k desde 0 a n.

Figura 1: Ejemplo de triángulo de pascal

Los números de la fila n cumplen que:
- El número combinatorio de la izquierda es $\binom{n}{0} = 0$
- El número combinatorio de la derecha es $\binom{n}{n} = 1$
- El número combinatorio $\binom{n}{k}$, con $k \neq 0$ y $k \neq n$, se obtiene de sumar los dos números combinatorios que están sobre él en la fila $(n - 1)$.

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

La utilidad del triángulo de Pascal es que entrega un algoritmo para encontrar de forma recursiva todos los números combinatorios, pero calculando sumas en lugar de multiplicaciones.

DISTRIBUCIÓN BINOMIAL

En experimentos donde uno está interesado en observar un número particular de “éxitos” en un número definido de “intentos”, se conocen como experimentos binomiales. Si definimos la variable aleatoria X cómo “el número de éxitos obtenidos en n intentos” y la

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

probabilidad de obtener un éxito es p en cada intento, entonces decimos que X se distribuye cómo una binomial de parámetros n y p, lo que denotamos como:

En un experimento Binomial, la probabilidad de obtener exactamente k éxitos en los n intentos (es decir, la probabilidad de que $X = k$), puede calcularse por medio de la siguiente fórmula:

$$\mathrm{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

En esta fórmula, el factor p^k corresponde a la ocurrencia de k éxitos y el factor $(1 - p)^{n-k}$ corresponde a los fracasos, que ocurren $n - k$ veces. El número combinatorio corresponde a que la ocurrencia de k éxitos y $(n - k)$ fracasos se puede dar exactamente de $\binom{n}{k}$ maneras.

RELACIONANDO BERNOULLI Y BINOMIAL

Si tenemos n variables aleatorias Bernoulli, W_1, W_2, ..., W_n, donde cada una de ellas tiene probabilidad de éxito p y es independiente de las otras variables, y definimos X como la suma de estas n variables de Bernoulli, entonces:

$$X = W_1 + W_2 + \cdots + W_n \sim \text{Binomial } (n, p)$$

Por ejemplo, supongamos que tenemos una moneda y definimos W_k como una variable aleatoria Bernoulli que toma el valor 1 si sale sello en el lanzamiento k y 0 si no, es decir, $W_k \sim Bernoulli(\frac{1}{2})$. Repetimos esto 10 veces y analizamos cuántos sellos obtuvimos. Luego, definimos X como la suma de las variables aleatorias asociadas a cada lanzamiento. Así, según lo que vimos anteriormente, X corresponde a una variable aleatoria Binomial, es decir, $X \sim Binomial(10, \frac{1}{2})$.

ESPERANZA Y VARIANZA DE UNA BINOMIAL

Considerando que una variable aleatoria que distribuyen como Binomial viene a partir de una suma de variables aleatorias Bernoulli W_1, W_2, ..., W_n.

$$X = W_1 + W_2 + \ldots + W_n \sim \text{Binomial } (n, p)$$

Por otro lado, sabemos que la esperanza de una suma de variables aleatorias, es la suma de las esperanzas de cada una de ellas. Utilizando que $E(W_i) = p$ para $i = 1, 2, 3, ..., n$ obtenemos que la esperanza de la variable aleatoria X es la siguiente:

$$\begin{aligned} \mathrm{E}(X) &= \mathrm{E}(W_1 + W_2 + W_3 + \ldots + W_n) \\ &= \mathrm{E}(W_1) + \mathrm{E}(W_2) + \mathrm{E}(W_3) + \cdots + \mathrm{E}(W_n) \\ &= \ p + p + p + \cdots + p \\ &= n \cdot p \end{aligned}$$

Por lo tanto, el valor esperado de una variable aleatoria X que distribuye como una Binomial está dada por:

$$\mathrm{E}(X) = n \cdot p$$

Observación

Si tenemos n variables aleatorias **independientes** W_i, y definimos una variable aleatoria $X = W_1 + W_2 + ... + W_n$, entonces la varianza de X puede calcularse de la siguiente forma:

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{Var}(W_1 + W_2 + W_3 + \ldots + W_n) \\ &= \mathrm{Var}(W_1) + \mathrm{Var}(W_2) + \mathrm{Var}(W_3) + \ldots + \mathrm{Var}(W_n) \end{aligned}$$

Luego, como W_1, W_2, ..., W_n son variables aleatorias **independientes**, podemos escribir la varianza de X de la siguiente forma:

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{Var}(W_1 + W_2 + \ldots + W_n) \\ &= \mathrm{Var}(W_1) + \mathrm{Var}(W_2) + \ldots + \mathrm{Var}(W_n) \end{aligned}$$

Por otro lado, podemos utilizar la siguiente expresión para la varianza:

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

Y considerar que $W_i = (W_i)^2$, ya que $W_i = 0$ o$W_i = 1$. Si reemplazamos esto en la igualdad anterior, obtenemos lo siguiente:

$$\begin{aligned} \mathrm{Var}(W_i) &= \mathrm{E}(W_i^2) - (\mathrm{E}(W_i))^2 \\ &= \mathrm{E}(W_i) - (\mathrm{E}(W_i))^2 \end{aligned}$$

Además, ya sabemos que $E(W_i) = p$. Reemplazando esto en la expresión anterior obtenemos lo siguiente:

$$\text{Var}(W_i) = p - p^2 = p \cdot (1 - p)$$

Finalmente, obtenemos lo siguiente:

$$\begin{aligned}\text{Var}(X) &= \text{Var}(W_1 + W_2 + \ldots + W_n) \\ &= \text{Var}(W_1) + \text{Var}(W_2) + \cdots + \text{Var}(W_n) \\ &= p \cdot (1-p) + p \cdot (1-p) + \cdots + p \cdot (1-p) \\ &= n \cdot p \cdot (1-p)\end{aligned}$$

Por lo tanto, la varianza de una variable aleatoria X que distribuye como una Binomial está dada por:

$$\text{Var}(X) = n \cdot p \cdot (1 - p)$$

SÍNTESIS

- Un experimento aleatorio es de tipo Bernoulli si sólo puede obtener dos resultados que se denominan como "éxito" y "fracaso". La variable aleatoria asociada a este experimento toma el valor 1 con probabilidad p, y el valor 0 con probabilidad $(1 - p)$.

- En términos generales, se dice que la variable aleatoria X se distribuye como una Bernoulli de parámetro p, con $0 \leq p \leq 1$, lo que se escribe de la siguiente manera:

$$X \sim Bernoulli(p)$$

- La función de probabilidad de la variable aleatoria X que se distribuye como una Bernoulli se puede expresar de la siguiente forma:

$$P(\text{Éxito}) = P(X = 1) = p$$
$$P(\text{Fracaso}) = P(X = 0) = (1 - p)$$

- Para una variable aleatoria X de tipo Bernoulli con parámetro p se tiene que:

- El valor esperado es igual a $E(X) = p$.
- La varianza es igual a $Var(X) = p \cdot (1 - p)$

- El número combinatorio $\binom{n}{k}$ se define como el número de combinaciones de tamaño k de un conjunto de $n$elementos, esto es, la cantidad de subconjuntos de tamaño k que pueden formarse a partir de extraer sin reposición elementos de un conjunto que son distinguibles entre sí.

- Hay dos expresiones usuales para calcular el número combinatorio, cada una basada en contar de manera distinta la cantidad de subconjuntos de tamaño k de un conjunto con n elementos.

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \qquad \binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

- El triángulo de Pascal es una representación en la que los números combinatorios se disponen en forma de triángulo, de tal forma que la fila n —ésima corresponde a los valores de los números $\binom{n}{k}$, con k desde 0 a n. En esta representación, un número combinatorio se calcula sumando los dos números combinatorios que están inmediatamente sobre él.

- Experimentos donde uno está interesado en observar un número particular de "éxitos" en un número definido de "intentos'' se conocen cómo experimentos binomiales.

- La variable aleatoria X definida cómo "el número de éxitos obtenidos en n intentos", donde cada intento es independiente de los otros y tiene probabilidad de éxito igual a p, se distribuye cómo una binomial con n y p. Esto se denota como:

$$X \sim \text{Binomial}(n, p)$$

- Si la variable aleatoria X se distribuye cómo una binomial con parámetros n y p entonces la probabilidad de obtener $X = k$ éxitos en n intentos puede calcularse por medio de la fórmula:

Curso: Probabilidad y estadística descriptiva
Unidad 3 : Situaciones o fenómenos que se modelan por medio de las distribuciones binomial y normal
Tema: Distribución binomial
Contenido: Coeficiente binomial y distribución binomial

$$\mathrm{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

- Una variable aleatoria X definida cómo la suma de n variables aleatorias de Bernoulli, las cuales todas tienen la misma probabilidad p de éxito y son todas independientes entre sí, se distribuye cómo una binomial con parámetros n y p.

- Un experimento aleatorio puede describirse con una variable aleatoria Binomial $X = W_1 + W_2 + ... + W_n$ si las variables aleatorias, $W_1, W_2, ..., W_n$, que se distribuyen como una Bernoulli son **independientes**.

- El valor esperado y la varianza de una variable aleatoria X que distribuye como una Binomial tienen un valor preciso que dependen del parámetro p, correspondiente a la probabilidad de éxito y de n que corresponde al número de intentos en los que se quiere observar el número de éxitos.
  - El valor esperado está dada por $E(X) = n \cdot p$.
  - La varianza está dada por $Var(X) = n \cdot p \cdot (1 - p)$