

APUNTES, MANUALES, PRESENTACIONES



Universidad de Jaén

Facultad de Humanidades y Ciencias de la Educación

Introducción a la teoría de la probabilidad. Tema 4 ADPI

Antonio José Ibáñez Molina
Yasmina Crespo Cobo

12/012025

Análisis de datos en psicología I



CREA

Tema 4. Introducción a la teoría de la probabilidad

- 1. Objetivo de la estadística inferencial y del cálculo de probabilidades**
- 2. Definiciones previas**
- 3. Definición de probabilidad**
 - 3.1 Definición clásica (o a priori)**
 - 3.2 Definición estadística, frecuentista (o a posteriori)**
- 4. Concepto de probabilidad condicional**
 - 2.1 Teorema de Bayes**
 - 2.2 Teorema de la adicción o suma**
 - 2.3 Teorema del producto**
- 5. Combinatoria**
- 6. Variables aleatorias**
- 7. Valor esperado y otros momentos**

1. Objetivo de la estadística inferencial y del cálculo de probabilidades

¿Por qué es importante la teoría de probabilidad?

El conocimiento obtenido a través de la experimentación se desarrolla considerando la incertidumbre inherente a los resultados. Esto implica analizar las probabilidades de que los datos puedan reproducirse o aplicarse en contextos más amplios, utilizando herramientas de cálculo probabilístico. En este sentido, el razonamiento basado en la probabilidad, que encuentra su estructura formal en la teoría de la probabilidad, constituye un elemento fundamental para la ciencia y el avance del conocimiento.

Específicamente:

- La inferencia estadística o generalización de los datos de una muestra a una población, está basada en cálculos probabilísticos. Esto es, la probabilidad de que las hipótesis que se plantean sean ciertas o falsas

En este tema veremos los aspectos más básicos de esta disciplina.

2. Definiciones previas

Experimento aleatorio, Espacio muestral y Suceso:

- **Experimento aleatorio:** Cualquier experimento que se puede repetir infinitas veces y cuyo resultado es incierto
- **Espacio muestral:** Todos los resultados posibles de un experimento aleatorio. Se denota como **E**
- **Suceso:** Es un subconjunto del espacio muestral. Se denota como **A**. Puede ser:

- **Elemental:** un solo resultado de E. Ej: que salga un 2 en un dado
- **Compuesto:** Varios resultados posibles. Ej: que salga un número par en un dado
- Los sucesos pueden ser también **seguros** si se dan siempre
- **Imposibles** si no se dan nunca
- **Incompatibles o excluyentes:** si no se pueden dar o verificar simultáneamente por no tener elementos comunes: *salir un número par o impar tras tirar un dado*
- **Complementario de un suceso:** Subconjunto de sucesos que no forman parte de ese suceso.
 - El complementario de sacar un 5 al lanzar el dado es: $\{1,2,3,4,6\}$
 - El complementario de sacar número par al lanzar el dado es: $\{1,3,5\}$
- **Independientes:** la ocurrencia de uno no afecta a la probabilidad de ocurrencia del otro (lo desarrollaremos un poco más después)

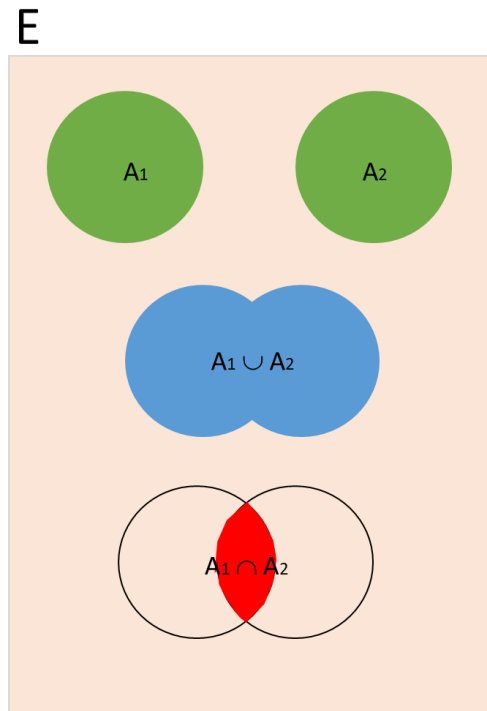
2.1 Operaciones con sucesos

Para sucesos A_1 y A_2 ...

- **Unión de sucesos** $A_1 \cup A_2$ son todos los resultados posibles en A_1 y A_2
- Ej: $A_1 = \{1, 2, 4\}$; $A_2 = \{2, 3, 6\}$; $A_1 \cup A_2 = \{1, 2, 3, 4, 6\}$

- **Intersección de sucesos** $A_1 \cap A_2$ son los resultados que pertenecen a A_1 y A_2 al mismo tiempo.
- Ej: $A_1 = \{1, 2, 4\}$; $A_2 = \{2, 3, 6\}$; $A_1 \cap A_2 = \{2\}$

En la siguiente figura, se representan los conceptos expuestos anteriormente utilizando diagramas o conjuntos de Venn.



3. Definición de probabilidad

La definición de probabilidad ha resultado problemática históricamente debido a que es un concepto que se entiende fácilmente mediante intuición pero matemáticamente es difícil de concretar. Nosotros consideraremos las dos definiciones más utilizadas.

3.1 Definición clásica (o a priori):

Razón entre **Frecuencia conocida** de un determinado evento o suceso y el total de casos posibles

¿Cuán frecuente es cierta característica 'x' en una población?

- Esto es: $P(x) = k/N$, donde k son los casos favorables (veces que sabemos que se da x) entre el total de casos posibles N (espacio muestral). **Siendo éstos igual de probables**

- Ejemplo: probabilidad de que salga 2 al lanzar un dado: $P(2)=1/6$

Problema de esta definición: Es una definición que depende "a priori" del conocimiento de ocurrencia de la probabilidad de cada evento o suceso. Por ejemplo, es problemático asignar de antemano que la probabilidad de que salga cara al tirar una moneda sea de 0,5 sólo porque existe una

posibilidad de un total de dos. ¿cómo estar seguros de que la moneda está bien equilibrada?

Lo mismo ocurre con la tirada de un dado ¿estamos seguros de que no existe un pequeño sesgo en el dado para ciertos valores? → no en todos los casos.

Además, otro problema es que necesitamos saber todos los casos posibles (N) del espacio muestral y eso en la práctica no siempre es posible.

3.2 Definición estadística, frecuentista (o a posteriori):

Esta forma de abordar este concepto parte de la idea de que la probabilidad es la frecuencia relativa de un determinado suceso.

Definición formal: Límite o valor al que tiende la frecuencia relativa cuando el número de ensayos (N) tiende a infinito que matemáticamente sería:

$$p(x) = \lim_{N \rightarrow \infty} \frac{n(x)}{N}$$

- Donde $n(x)$ es la frecuencia absoluta y N el numero de veces que se realiza el experimento. Es decir, $n(x)/N$ sería la frecuencia relativa para N experimentos
- Ejemplo: si voy lanzando un dado muchas veces y cuantificando las veces que sale 2, me daré cuenta de que cuantas más veces realizo este experimento, la frecuencia relativa tiende a un número concreto (que será 1/6 si el dado no está sesgado).
- Observa que aquí N no se refiere al espacio muestral (en el caso del dado sería 6) sino a todos los ensayos o intentos de un experimento dado

La razón por la que se denomina definición *a posteriori* es porque la probabilidad se calcula a partir de los resultados del experimento aleatorio

En la siguiente gráfica hemos incluido una simulación de las veces que sale cara al tirar una moneda (probabilidad simulada de 0,5). En el eje de la Y se representa la frecuencia relativa de cara ($n(\text{cara})/N$) y en el eje de la X, se representa N. Como se puede observar, conforme aumenta N, la tendencia de la probabilidad es hacia el valor 0,5.



4. Concepto de probabilidad condicional

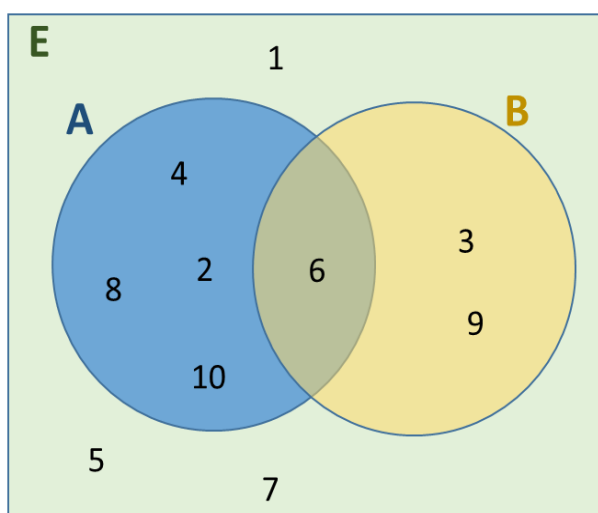
Se refiere a la probabilidad de un suceso A teniendo en cuenta que ha ocurrido otro suceso B. Se denota como $P(A|B)$, donde la línea vertical podría leerse como "dado el suceso". Así:

$P(A|B)$ = "probabilidad de A dado que ya ha sucedido B"

Pues bien, la probabilidad condicional cumple la siguiente relación:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Esta relación no requiere una demostración muy compleja. Basta con que consideres el siguiente ejemplo y medites sobre ella tal como hicimos en clase:



Experimento aleatorio con $E = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$; es decir, cada vez que se realiza el experimento saldría uno de esos números. Suponemos que son equiprobables

Definimos también:

$A = \{2, 4, 6, 8, 10\}$ y $B = \{3, 6, 9\}$.

Y decimos que ha ocurrido A o B cada vez que ocurre alguno de los números que lo componen.

$P(A|B)$ sería la probabilidad de que ocurriera 6, dado que ha ocurrido B (que puede ser 3, 6 ó 9). Es decir $1/3$

$P(A \cap B)$ sería $1/10$

$P(B)$ sería $3/10$

Y es evidente que $\frac{P(A \cap B)}{P(B)} = 1/3$

4.1 Teorema de Bayes

Se trata de una relación importante en el razonamiento probabilístico y que se deriva directamente de la probabilidad condicional:

Si consideramos que

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ y que $P(B|A) = \frac{P(A \cap B)}{P(A)}$ si despejamos en ambas expresiones $P(A \cap B)$ e igualamos los resultados obtenemos:

$$p(A|B) = \frac{p(B|A) p(A)}{p(B)}$$

Que es una igualdad muy apreciada porque relaciona las probabilidades condicionadas contrarias para dos sucesos. En muchas ocasiones conocemos por ejemplo $p(B|A)$

Y quisiéramos saber $p(A|B)$. Si eso ocurre, es posible que sea más fácil obtener $p(A)$ y $p(B)$, y utilizar la fórmula de Bayes, que hacer un estudio para hallar $p(A|B)$.

4.2 Teorema de la adición o de la suma

Este teorema de fácil comprobación nos indica que la probabilidad de que alguno de los dos sucesos (o los dos a la vez) es la suma de las probabilidades de cada suceso menos la probabilidad de su intersección. Fíjate que $A \cup B$ se puede pensar como "A ó B" en el sentido lógico de que podrían darse cualquiera.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Ejemplo: Tras realizar una encuesta encontramos que el 70% de los encuestados usa transporte público, el 40% usa su propio coche y el 30% usa ambos tipos de transporte. Calcula la probabilidad de que se use algún tipo de transporte?

Para calcular la probabilidad de que una persona use algún tipo de transporte, necesitamos considerar que el espacio total está compuesto por aquellos que usan transporte público, coche propio, o ambos (intersección). Usaremos la fórmula de la unión de conjuntos para resolverlo:

$$P(\text{Usar algún transporte}) = P(TP) + P(C) - P(TP \cap C)$$

Donde:

- $P(TP) = 0,7$ (probabilidad de usar transporte público)
- $P(C) = 0.4$ (probabilidad de usar coche propio)
- $P(TP \cap C) = 0.3$ (probabilidad de usar ambos tipos de transporte).

Sustituyendo los valores:

$$P(\text{Usar algún transporte}) = 0.7 + 0.4 - 0.3 = 0,8$$

Por lo tanto, la probabilidad de que se use algún tipo de transporte es 80%

4.3 Teorema del producto

Para dos sucesos A y B independientes se verifica que:

$$P(A \cap B) = P(A) \cdot P(B)$$

Que si lo pensamos bien se deriva del hecho de que si A y B son independientes, entonces $p(A|B) = p(A)$ ya que por definición, el hecho de que ocurra B no afecta a la ocurrencia de A.

Fíjate que $A \cap B$ se puede pensar como "A y B" en el sentido lógico de que se dan los dos a la vez.

Este teorema es enormemente útil porque nos va a ayudar a calcular la probabilidad de que ocurran en un mismo momento dos eventos que no tienen relación.

Ejemplo. En una clase con 15 chicos y 12 chicas, hay 10 chicos y 8 chicas que han aprobado Análisis de Datos. Eligiendo un alumno al azar, calcula la probabilidad de que sea chico y no apruebe Análisis. (*planteamos el enunciado con un diagrama de árbol*)

$$P(A \cap B) = P(A) \cdot P(B/A) = 15/27 * 5/15 = 0,18$$

5. Combinatoria

Se estudian las diferentes formas de agrupar y combinar un conjunto de elementos.

Ejemplos:

- ¿De cuantas formas podéis sentaros en los distintos asientos de esta clase?
- ¿Con 100 preguntas, cuántos exámenes de 25 preguntas puedo construir?

Para resolver este tipo de preguntas se estudian las variaciones, permutaciones y combinaciones

- **Partimos de un principio fundamental:** si un experimento se puede hacer de n_1 maneras, y a su vez cada manera de n_2 maneras, etc. Las formas totales en las que puede hacerse el experimento es de $n_1 \times n_2 \dots$
 - Ejemplo: ¿De cuántas maneras podemos hacer la ruta Linares → Jaén → Andújar → Córdoba, si de Linares a Jaén hay 3 rutas, de Jaén a Andújar 2 rutas y de Andújar a Córdoba hay 4 rutas?

Solución: $3 \times 2 \times 4 = 24$ rutas posibles

En esta sección estudiaremos las permutaciones, variaciones y combinaciones SIN reposición.

Permutaciones: Distintas agrupaciones de m elementos

Las permutaciones son las distintas maneras que tenemos de ordenar un conjunto de m elementos. Por tanto, las agrupaciones **difieren en el orden** de los elementos

Se calculan como:

- $P_m = m! = m \cdot (m - 1) \cdot \dots \cdot 2 \cdot 1$

El símbolo “!” indica el factorial del número que lo precede.

- Por convención y conveniencia: $0! = 1$

Ejemplo: ¿De cuantas formas pueden intercambiarse en sus asientos las personas de la primera fila en una clase?

Variaciones: Es un caso más complejo porque **las agrupaciones se hacen con un número menor al total** $n < m$.

- El orden de los elementos determina también las agrupaciones. Entonces, las variaciones son sobre m elementos tomados de n en n (de dos en dos, tres en tres etc)
- En un juego de roles, ¿cuántas parejas jefe-empleado podemos hacer con las personas que hay en una clase? Considera que en este caso las personas A-B y B-A serían parejas distintas puesto que cambian sus roles.

$$V_{m,n} = \frac{m!}{(m - n)!}$$

Que se lee: variaciones de un conjunto m tomadas de n en n

- En el ejemplo del número de parejas jefe-empleado, si hubiese 10 personas en la clase, el número de parejas posibles sería:

$$V_{10,2} = \frac{10!}{(10-2)!} = \frac{10 \cdot 9 \cdot 8!}{8!} = 90$$

Combinaciones: Las agrupaciones se realizan con un número menor del total $n < m$, pero en este caso no importa el orden. Por tanto, las agrupaciones se distinguen por los elementos que la componen y no por su orden.

Ejemplo: ¿Cuántas parejas podemos hacer con las personas que hay en clase? Fíjate que en este caso, la pareja A-B sí sería la misma que B-A.

$$C_{m,n} = \frac{m!}{n! \cdot (m-n)!} = \frac{V_{m,n}}{n!}$$

Que se lee: combinaciones de un conjunto m tomado de n en n.

Solución para el número de parejas en el caso de haber 10 alumnos en clase: $C_{10,2} = \frac{10!}{2! \cdot (10-2)!} = \frac{10 \cdot 9 \cdot 8!}{2! \cdot 8!} = 45$

En este caso, es muy interesante saber que

$$C_{m,n} = \binom{m}{n} \quad \text{siendo esta última expresión el coeficiente binomial}$$

El coeficiente binomial no solo se usa en combinatoria, sino que también se utiliza en otros campos, y es fundamental en el desarrollo de la expresión $(x+y)^n$. El teorema de Newton establece la siguiente relación:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

que es muy útil para calcular las potencias de las sumas (este teorema excede los objetivos del curso y no se evaluará).

Además, una bonita manera de saber los valores de los coeficientes binomiales es a través del triángulo de Pascal o pirámide de Tartaglia. Se trata de un triángulo de números que se construye poniendo un 1 en el vértice superior y construyendo hacia abajo una pirámide de números que son la suma de los dos inmediatamente superiores (ver la pirámide derecha de la figura inferior). Pues bien, esta pirámide son los resultados de los coeficientes binomiales ordenados en otra pirámide tal como aparece en la pirámide izquierda de la figura inferior (esta parte es una curiosidad que tampoco se evaluará en este curso).

$$\begin{array}{c}
 \binom{0}{0} \\
 \binom{1}{0} \quad \binom{1}{1} \\
 \binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2} \\
 \binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3} \\
 \text{etc}
 \end{array}
 =
 \begin{array}{c}
 1 \\
 1 \quad 1 \\
 1 \quad 2 \quad 1 \\
 1 \quad 3 \quad 3 \quad 1 \\
 1 \quad 4 \quad 6 \quad 4 \quad 1 \\
 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
 \text{etc}
 \end{array}$$

6. Las variables aleatorias

Ya sabemos que una variable aleatoria es el resultado de un experimento aleatorio. En esta sección vamos a profundizar específicamente en ellas por su relevancia en la estadística.

Definición:

Una Variable Aleatoria (VA) es una función en la que **se asigna un valor**, y sólo uno, **a cada suceso posible de un espacio muestral**, y se asigna **una probabilidad** a cada uno de los valores. Esto implica construir un modelo de distribución de probabilidad.

Ejemplo, El experimento aleatorio "tirar un dado" produce una variable aleatoria con $E = \{1,2,3,4,5,6\}$. Cada x tendrá una probabilidad $p(x) = 1/6$

Las variables aleatorias no son empíricas sino teóricas. Conocemos las probabilidades de los sucesos o eventos. Por el contrario, en una variable empírica es posible que no conozcamos todos los valores posibles

- De modo similar a como vimos en temas anteriores, **las VA pueden ser discretas o continuas**
 - Discretas: La variable se define sobre un espacio muestral numerable (entre dos valores consecutivos no hay otros valores posibles). Se usan números enteros
 - Continuas: Las variables se definen sobre un espacio muestral no numerable (entre dos valores cualquiera, existen infinitos valores posibles). Se usan números decimales

Ejemplos de posibles variables aleatorias discretas y continuas:

- Género de una persona seleccionada al azar en nuestro grupo de clase (discreta)
- Tiempo de reacción ante la pregunta "¿cuántos años tienes?" de una persona tomada al azar en nuestro grupo de clase.
- Número de ojos de una mosca de la fruta (discreta)

Distribución de probabilidad de variables discretas

Se lleva a cabo con la función de probabilidad o mediante la función de distribución. Ambas son funciones matemáticas que establecen una relación entre valores posibles y su probabilidad.

Función de probabilidad:

Se trata de aquella función que relaciona cada valor de la variable con su probabilidad de ocurrencia. A esta probabilidad se le suele llamar también probabilidad puntual. El modo de escribirla y un ejemplo lo ponemos en la figura siguiente.

$$f(x_i) = P(x = x_i)$$

x	f(x)
0	0,25
1	0,5
2	0,25

Función de probabilidad

Función de distribución:

Se trata de aquella función que relaciona cada valor de la variable con su probabilidad acumulada de ocurrencia. Esta probabilidad es la que se refiere a la probabilidad de obtener cualquier valor menor o igual que x, que se escribe $P(X \leq x_i)$ y que se corresponde con la suma de las probabilidades puntuales de esos valores. Evidentemente, la suma de todas las probabilidades puntuales será 1. Ver la figura siguiente para la notación y un ejemplo.

$$F(x_i) = P(x \leq x_i) = \sum_{i=1}^n f(x_i) = 1$$

x	F(x)
0	0,25
1	0,75
2	1

Función de distribución (acumulada)

Distribución de probabilidad de variables continuas

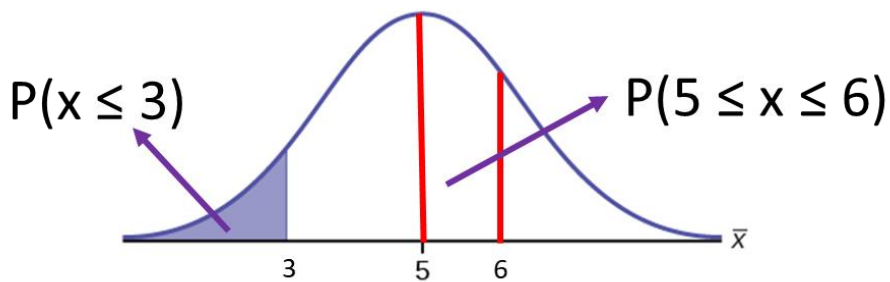
En este caso, la variable aleatoria se caracteriza por una función de densidad de probabilidad. Es el mismo concepto que hemos usado para las variables discretas, pero se le llama función de densidad porque, al ser continua, está densamente distribuida.

Cuando trabajamos con variables continuas, lo que nos interesará en la práctica es plantearnos preguntas sobre la probabilidad de que ocurran cierto grupo de valores

Ejemplos:

- Probabilidad de que la variable aleatoria tome un valor mayor de 140 que se escribiría: $P(x \geq 140)$
- Probabilidad de que la VA tome un valor entre 140 y 160 que se escribiría $P(140 \leq x \leq 160)$

Y gráficamente para una variable aleatoria con función de densidad de probabilidad igual a la normal, sería:



7. Momento esperado y otros momentos

Del mismo modo que se han analizado las propiedades de las distribuciones de frecuencia de las variables estadísticas (tendencia central, variabilidad, sesgo y forma), se pueden analizar las distribuciones de probabilidad de las variables aleatorias mediante **una serie de índices llamados momentos**:

- **Momento valor esperado o esperanza matemática** (E) ó μ
 - Análogo a la *media* de variables estadísticas
 - Definición: Sería la media que se obtendría en caso de que se observase un número infinito de valores de la variable aleatoria

Se calcula como

$$\mu_1 = E(X) = \sum_{i=1}^n x_i \cdot f(x_i) = \sum_{i=1}^n x_i \cdot P(X = x_i)$$

Es decir, multiplicando cada valor por su probabilidad puntual y sumándolos

Ejemplo:

x	f(x)
0	0,25
1	0,5
2	0,25

$$\mu = (0 \cdot 0,25) + (1 \cdot 0,5) + (2 \cdot 0,25) = 1$$

- Momento **Varianza σ^2 de una variable aleatoria discreta**
 - Análogo a la varianza de variables estadísticas
 - Definición: es la varianza que se obtendría sobre los valores observados en caso de que el número de observaciones creciera infinitamente

Se calcula como:

$$\sigma^2 = \sum_{i=1}^n (x_i - \mu)^2 \cdot f(x_i)$$

Que como se puede comprobar es una expresión muy similar a la de la varianza para un conjunto de datos empíricos. La diferencia está en sustituir $1/N$ por el valor de la probabilidad.

Ejemplo:

x	f(x)
0	0,25
1	0,5
2	0,25

$$\sigma^2 = [(0-1)^2 \cdot 0,25] + [(1-1)^2 \cdot 0,5] + [(2-1)^2 \cdot 0,25] = 0,5$$

Con lo visto hasta aquí sobre los momentos, ¿sabrías resolver el siguiente ejercicio?

- La probabilidad de que cierta máquina tragaperras dé un premio de 300 euros es de 0,002; de que dé 20 euros es de 0,01; y en el resto de jugadas no da nada. Teniendo en cuenta que cada partida vale 1 euro, ¿cuál es la esperanza de esta variable? ¿y la varianza?