Probabilidad y Estadística Fundamental

Principios de variables aleatorias

Profesor: Nicolás López

Universidad Nacional de Colombia

Contenido

Contenido

Introducción

Recordemos algunas definiciones

1. Variable
2. Experimento aleatorio.
3. Espacio y espacio muestral.

Introducción

Al realizar un experimento aleatorio, Ω puede ser arbitrario ó tedioso de describir:

- Al lanzar una moneda $\Omega = \{C, S\}$ con C representando cara y S sello, ó quizá $\Omega = \{H, T\}$ con H representando cara y T sello.
- Al buscar el primer enfermo de un conjunto de personas $\Omega = \{(E), (S, E), (S, S, E), (S, S, S, E)...\}$ siendo E la persona es enferma y S, sana.

Introducción

Por otra parte, puede que no estemos interesados en los resultados del experimento aleatorio en si mismos, sino en ciertos **valores numéricos** a partir de ellos:

- Al lanzar dos monedas una vez $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$ puedo estar interesado en el número de caras obtenidas (0, 1 ó 2), más que en los $\omega \in \Omega$.
- Si nos interesan los ingresos de los servidores públicos de Colombia, al seleccionar uno de sus miembros, estamos interesados realmente en conocer su salario.
- ...

Introducción

Quizá sea buena idea hacer un **cambio de espacio**, en el cual sea menos difícil estudiar las propiedades de interés del experimento aleatorio:

- **Transportar** Ω a Ω', dónde Ω' es más fácil de estudiar.

Un espacio numérico que **estandarice** las propiedades de experimentos aleatorios similares.

Introducción

Ejemplo. Lanzamiento de una moneda una vez

Experimento aleatorio Lanzamiento de una moneda una vez.
Espacio muestral (de **salida**) $\Omega = \{C, S\}$.
Regla de asignación Número de caras observadas.
Espacio muestral (de **llegada**) $\Omega' = \{1, 0\}$.

Ejemplo. Lanzamiento de una moneda dos veces

Experimento aleatorio Lanzamiento de una monedas dos veces.
Espacio muestral (de **salida**) $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$.
Regla de asignación Número de caras observadas.
Espacio muestral (de **llegada**) $\Omega' = \{2, 1, 0\}$.

La idea es estudiar las características del experimento aleatorio ahora desde este nuevo espacio Ω', el cual es más sencillo de manipular.

Contenido

Variable aleatoria

Una regla (o también llamada *función*) X la cual está determinada por el resultado del experimento aleatorio, es decir, asigna a cada elemento $\omega \in \Omega$ un elemento $X(\omega) = x \in \Omega'$.

Ejemplo. Lanzamiento de una moneda dos veces

Experimento aleatorio Lanzamiento una moneda dos veces.
Espacio muestral $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$.
Variable aleatoria $X=$ *Número de caras observadas*.
Valores de la v.a

$$X((C, C)) = 2$$

$$X((C, S)) = 1$$

$$X((S, C)) = 1$$

$$X((S, S)) = 0$$

.

Variable aleatoria

Ejemplos

Algunos otros ejemplos de variables aleatorias (v.a.) son

1. $W =$ *Número de defectos observados en el proceso de fabricación*.
2. $Y =$ *Número de cestas realizadas en un partido de baloncesto*.
3. $Z =$ *Calificaciones de los estudiantes del curso de Estadística*.
4. $T =$ *Tiempo de espera en línea para la toma de un pedido*.

Ejemplo. Lanzamiento de una moneda dos veces

Volviendo al lanzamiento de una moneda dos veces, es claro que si la probabilidad de obtener cara es $p \in [0, 1]$, el vector de probabilidades en el espacio de salida Ω está dado por:

$$
\begin{aligned}
P(E_1) &= P(\{(C, C)\}) = p * p \\
P(E_2) &= P(\{(C, S)\}) = p * (1 - p) \\
P(E_3) &= P(\{(S, C)\}) = (1 - p) * p \\
P(E_4) &= P(\{(S, S)\}) = (1 - p) * (1 - p)
\end{aligned}
$$

Conformado por 4 eventos simples.

Ejemplo. Lanzamiento de una moneda dos veces

Por otra parte, el vector de probabilidades para el espacio de llegada Ω' está dado por:

$$
\begin{aligned}
P(X=2) &= P(\{\omega \in \Omega : X(\omega)=2\}) = P(\{(C,C)\}) = p * p \\
P(X=1) &= P(\{\omega \in \Omega : X(\omega)=2\}) = P(\{(C,S),(S,C)\}) \\
&= p*(1-p)+(1-p)*p = 2p*(1-p) \\
P(X=0) &= P(\{\omega \in \Omega : X(\omega)=2\}) = P(\{(C,C)\}) = (1-p)*(1-p)
\end{aligned}
$$

Conformado por 3 eventos simples. En general asumimos que $\Omega' = \mathrm{R}$, con lo cual la variable X del ejemplo es una variable aleatoria **real**.

Contenido

Cuantificación del experimento aleatorio

Introducción

Ahora que tenemos los valores numéricos a partir del experimento aleatorio transportado a **real**, **hemos formalizado la definición de una variable estadística**. También podemos formalizar características como:

- Distribución de la v.a.
- Promedio de la v.a.
- Varianza de la v.a.
- ...

Para ello diferenciamos dos tipos de variables aleatorias reales:

- **Discretas**: Toman un conjunto discreto de valores en $\mathbb{R}$.
- **Continuas**: Toman un conjunto continuo de valores en $\mathbb{R}$.

Ejercicio. Lanzamiento de una moneda tres veces

La misma moneda de los ejemplos anteriores es lanzada ahora tres veces:

1. ¿A qué es igual es espacio muestral del experimento?
2. ¿Cuántos elementos tiene el espacio muestral del experimento?
3. Si nos interesa el número de caras obtenidas en los lanzamientos ¿cómo debe definirse la v.a real correspondiente?
4. ¿Es la v.a. definida en el punto anterior discreta o continua?
5. ¿A qué es igual el vector de probabilidades del espacio de llegada $\mathbb{R}$?

Ejercicio. Lanzamiento de una moneda tres veces

1. ¿A qué es igual es espacio muestral del experimento?. Por comprensión: $\Omega = \{(r_1, r_2, r_3) : r_i \in \{C, S\}, i = 1, 2, 3\}$
2. ¿Cuántos elementos tiene el espacio muestral del experimento?. Por la regla fundamental del conteo extendida, $|\Omega| = 2 \times 2 \times 2 = 8$.
3. Si nos interesa el número de caras obtenidas en los lanzamientos ¿cómo debe definirse la v.a real correspondiente?. En este caso, $X =$ *Número de caras obtenidas en la ejecución del experimento*.
4. ¿Es la v.a. definida en el punto anterior discreta o continua?. La variable es discreta, pues toma los valores en $\{0, 1, 2, 3\}$, los cuales son una colección discreta de valores en $\mathbb{R}$.

Ejercicio. Lanzamiento de una moneda tres veces

5 ¿A qué es igual el vector de probabilidades del espacio de llegada R?. Este vector se puede resumir de la siguiente forma:

$$P(X=x)=\begin{cases}(1-p)*(1-p)*(1-p) & \text{si } x=0\\ 3*p*(1-p)*(1-p) & \text{si } x=1\\ 3*p*p*(1-p) & \text{si } x=2\\ p*p*p & \text{si } x=3\end{cases}$$

Variables aleatorias discretas

Función másica de probabilidad

La **función másica de probabilidad** (fmp) para una variable aleatoria real **discreta** X con valores $x_1, \ldots$ está dada por

$$f_X(x) = \begin{cases} P(X = x) & \text{si } x = x_1, \ldots \\ 0 & \text{en otro caso (eoc)} \end{cases}$$

Se destaca que fuera del rango de la función X, la probabilidad es siempre 0.

La **función de distribución** para una variable aleatoria real X está dada por

$$F_X(x) = P(X \leq x) \text{ con } x \in \mathrm{R}$$

Se destaca que la función está definida sobre todos los reales, independientemente de si la variable es discreta o continua.

Ejercicio. Lanzamiento de una moneda tres vez

5. ¿A qué es igual la función de distribución de la variable aleatoria real discreta X? Dibújela para $p = 0{,}5$.

Ejercicio. Lanzamiento de una moneda tres veces

5. ¿A qué es igual la función de distribución de la variable aleatoria real discreta X?. En este caso se tiene que los valores posibles de la variable definen los *saltos* que tendrá la función de distribución:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ (1-p)*(1-p)*(1-p) & \text{si } x \geq 0 \text{ y } x < 1 \\ (1-p)*(1-p)*(1-p) + 3*p*(1-p)*(1-p) & \text{si } x \geq 1 \text{ y } x < 2 \\ (1-p)*(1-p)*(1-p) + 6*p*p*(1-p) & \text{si } x \geq 2 \text{ y } x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}$$

Note que para cualquier $x \in \mathrm{R}$ puede calcular $F_X(x)$ aún siendo X discreta.

¿A qué es igual $F_X(x)$ si $p = 0{,}5$?

Variables aleatorias discretas

Valor esperado (media)

Para una v.a real, de manera análoga a la media aritmética del escenario descriptivo, podemos calcular un valor que represente dicha variable. Este valor corresponde al *valor esperado* o promedio de la v.a. Para una v.a discreta está dado por:

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

Es un promedio (ponderado por la probabilidad de ocurrencia) de los valores que toma la v.a.

Ejercicio. Distribución de número de hijos

Una trabajadora social encuentra que para una población determinada, la función másica de probabilidad para la v.a. X, que denota el número de hijos, está dada por:

$$f_X(x) = \begin{cases} 0{,}10 & \text{si } x = 0 \\ 0{,}40 & \text{si } x = 1 \\ 0{,}20 & \text{si } x = 2 \\ 0{,}15 & \text{si } x = 3 \\ 0{,}10 & \text{si } x = 4 \\ 0{,}05 & \text{si } x = 5 \\ 0. & \text{en otro caso} \end{cases}$$

¿Cuántos hijos se espera que tenga una mujer en esta población?

Ejercicio. Distribución de número de hijos

Se pregunta el valor esperado de la v.a. X, por lo cual, al ser v.a. discreta, se tiene:

$$E(X) = \sum_{i=1}^{5} x_i P(X = x_i) = 0 \times P(X = 0) + ... + 5 \times P(X = 5) = 1{,}90$$

Se espera que una mujer en esta población tenga en promedio 1,90 hijos.

Variables aleatorias discretas
Varianza

Para una v.a real, de manera análoga a la varianza del escenario descriptivo, podemos calcular un valor que represente la variablidad de los datos respecto al valor esperado. Este valor corresponde a la *varianza* de la v.a. Para una v.a discreta está dado por:

$$V(X) = \sum_{i=1}^{n} (x_i - E(X))^2 P(X = x_i)$$

Es un promedio (ponderado por la probabilidad de ocurrencia) de los desvíos al cuadrado respecto al valor central.

Ejercicio. Distribución de número de hijos

Con su ayuda, la trabajadora social encontró que una mujer de la población de estudio tiene en promedio 1.9 hijos. Ella se pregunta ahora por la variabilidad de la v.a subyacente, para lo cual, desea calcular una medida resumen de variabilidad para el fenómeno aleatorio. ¿A qué es igual el coeficiente de variación de la variable aleatoria X?

Ejercicio. Distribución de número de hijos

Se pregunta el coeficiente de variación de la v.a. X, para lo cual se necesita primero calcular la varianza de la variable. Al ser v.a. discreta, se tiene:

$$\begin{aligned} V(X) &= \sum_{i=1}^{5}(x_i - E(X))^2 P(X = x_i) \\ &= (0 - 1{,}9)^2 P(X = 0) + ... + (5 - 1{,}9)^2 P(X = 5) = 1{,}79 \end{aligned}$$

Con lo cual, la varianza de X es 1,79 hijos^2, o de manera equivalente, X tiene una desviación estándar de 1,34 hijos. Además:

$$CV(X) = \frac{sd(X)}{E(X)} = \frac{1{,}34}{1{,}9} = 0{,}70$$

Así, la variabilidad respecto a la media para la v.a X es bastante alta, pues $CV(X) = 0{,}70$.