Introducción a la probabilidad y estadística CM274

César Lara Avila

14 de septiembre de 2017

https://github.com/C-Lara

5b. Varianza de variables aleatorias discretas

Difusión

El valor esperado (media) de una variable aleatoria es una medida de localización o tendencia central. Si tuviera que resumir una variable aleatoria con un solo número, la media sería una buena opción.

Sin embargo, la media deja fuera una buena cantidad de información. Por ejemplo, las variables aleatorias X e Y tienen la media 0, pero su masa de la probabilidad se separa de la media totalmente diferentemente.

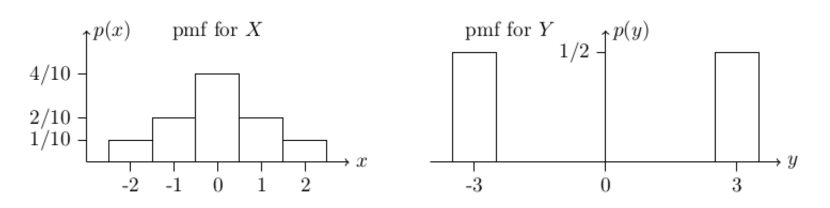
valores x	-2	-1	0	1	2
pmf p(x)	1/10	2/10	4/10	2/10	1/10

valores y	-3	3
pmf p(y)	1/2	1/2

Difusión

Probablemente sea un poco más fácil ver los diferentes propagaciones en las gráficas de las funciones de masa de probabilidad. Utilizamos barras en lugar de puntos para dar una mejor percepción de la masa.

El gráfico siguiente es la función de masa de probabilidad para dos distribuciones con media 0.



Varianza y desviación estándar

Tomando la media como el centro de la distribución de probabilidad de una variable aleatoria, la varianza es una medida de cuanto la masa de probabilidad se extiende alrededor de este centro. Empezaremos con la definición formal de la varianza y luego revisaremos su significado.

Si X es una variable aleatoria con media $\mathbb{E}(X) = \mu$, entonces la varianza de X se define por:

$$\text{Var}(X) = \mathbb{E}((X - \mu)^2).$$

La desviación estándar σ de X es definida por

$$\sigma = \sqrt{\text{Var}(X)}.$$

Varianza y desviación estándar

Si la variable aleatoria relevante es clara desde el contexto, entonces la varianza y la desviacin estándar se denotan a menudo por σ^2 y σ , así como la media es μ.

¿Qué significa esto? Primero, vamos a reescribir la definición explícitamente como una suma. Si X toma valores $x_1, x_2, \ldots, x_n$ con función de masa de probabilidad $p(x_i)$ entonces

$$\mathrm{Var}(X) = \mathbb{E}((X - \mu)^2) = \sum_{i=1}^{n} p(x_i)(x_i - \mu)^2.$$

En palabras, la fórmula de $\mathrm{Var}(X)$ dice que se toma un promedio ponderado de la distancia cuadrada a la media. Mediante la cuadratura, nos aseguramos de que estamos promediando sólo valores positivos, de modo que la propagación a la derecha de la media no cancele la de la izquierda.

Varianza y desviación estándar

Usando esperanza, estamos ponderando valores de probabilidad altos más que valores de probabilidad bajos.

Notas sobre unidades:

1. σ tiene las mismas unidades que X.
2. $\text{Var}(X)$ tiene las mismas unidades que el cuadrado de X. Así que si X está en metros, entonces $\text{Var}(X)$ está en metros al cuadrado. Debido a que σ y X tienen las mismas unidades, la desviación estándar es una medida natural de propagación.

Ejemplos

1. Para cada variable X, Y, Z y W dibujamos la función de masa de probabilidad y calculamos la media y la varianza.

Función de masa de probabilidad X

valores x	1	2	3	4	5
pmf p(x)	1/5	1/5	1/5	1/5	1/5

Función de masa de probabilidad Y

valores y	1	2	3	4	5
pmf p(y)	1/10	2/10	4/10	2/10	1/10

Función de masa de probabilidad Z

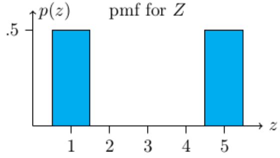
valores z	1	2	3	4	5
pmf p(z)	5/10	0	0	0	5/10

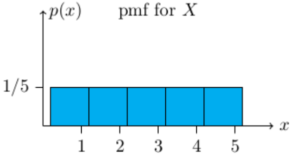
Ejemplos

Función de masa de probabilidad W

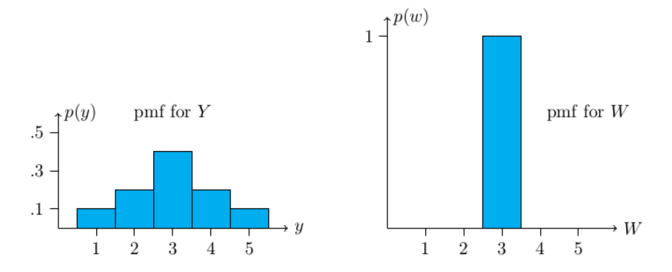
valores w	1	2	3	4	5
pmf p(w)	0	0	0	0	0

Cada variable aleatoria tiene la misma media 3, pero la probabilidad se difunde de manera diferente. En los diagramas a continuación, mostramos los pmf de mayor a menor varianza: Z, X, Y, W.





Ejemplos



A continuación verificamos nuestra intuición visual calculando la varianza de cada una de las variables. Todos ellos tienen como media $\mu = 3$. Dado que la varianza se define como un valor esperado, podemos calcularlo usando tablas.

valores x	1	2	3	4	5
pmf p(x)	1/5	1/5	1/5	1/5	1/5
(X - μ^2)	4	1	0	1	4

Ejemplos

$$\text{Var}(X) = \mathbb{E}((X-\mu)^2) = \frac{4}{5} + \frac{1}{5} + \frac{0}{5} + \frac{1}{5} + \frac{4}{5} = 2$$

valores y	1	2	3	4	5
pmf p(y)	1/10	2/10	4/10	2/10	1/10
(Y - μ^2)	4	1	0	1	4

$$\text{Var}(Y) = \mathbb{E}((Y-\mu)^2) = \frac{4}{10} + \frac{2}{10} + \frac{0}{10} + \frac{2}{10} + \frac{4}{10} = 1.2$$

valores z	1	2	3	4	5
pmf p(z)	5/10	0	0	0	5/10
(Z -μ^2)	4	1	0	1	4

$$\text{Var}(Z) = \mathbb{E}((Z-\mu)^2) = \frac{20}{10} + \frac{20}{10} = 4.$$

Ejemplos

valores w	1	2	3	4	5
pmf p(w)	0	0	0	0	0
(W -μ^2)	4	1	0	1	4

$$\text{Var(W)} = 0.$$

Ten en cuenta que W no varía, por lo que tiene varianza 0!.

2. Si $X \sim \text{Bernoulli}(p)$ entonces $\text{Var}(X) = p(1-p)$.

En efecto, sabemos que $\mathbb{E}(X) = p$. Ahora calculamos $\text{Var}(X)$ usando una tabla.

valores X	0	1
pmf p(x)	1 -p	p
(X -μ^2)	$(0-p)^2$	$(1-p)^2$

Ejemplos

$$\mathrm{Var}(X) = (1-p)p^2 + p(1-p)^2 = (1-p)p(1-p+p) = (1-p)p$$

Como con todas las cosas de Bernoulli, debe recordar esta fórmula.

Pregunta: ¿ Para qué valor de p tiene Bernoulli(p) tiene la varianza más alta?.

Independencia

Hasta ahora hemos estado utilizando la noción de variable aleatoria independiente sin nunca cuidadosamente definirla. Por ejemplo, una distribución binomial es la suma de ensayos independientes de Bernoulli.

En un sentido probabilístico las variables aleatorias X e Y son independientes si conociendo el valor de X no da ninguna información sobre el valor de Y.

Por ahora podemos usar la siguiente definición, que es válida para variables aleatorias discretas.

Definición:

Las variables aleatorias discretas X e Y son independiente si

$$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b)$$

para valores a, b. Es decir, las probabilidades se multiplican.

Propiedades de la varianza

Las tres propiedades más útiles para calcular la varianza son:

1. Si X y Y son independientes entonces $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
2. Para constantes a y b, $\text{Var}(aX + b) = a^2\text{Var}(X)$.
3. $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Para la propiedad 1, anote cuidadosamente el requisito de que X e Y son independientes.

La propiedad 3 da una fórmula para $\text{Var}(X)$ que a menudo es más fácil de usar en cálculos manuales.

Ejemplos

1. Usa la propiedad 3 para calcular la varianza de $X \sim \text{Bernoulli}(p)$.

Desde la tabla,

valores X	0	1
pmf p(x)	1 -p	p
X^2	0	1

tenemos que $\mathbb{E}(X^2) = p$. Así con la propiedad 3 produce,

$$\text{Var(X)} = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1-p).$$

2. Supongamos $X \sim \text{Binomial}(n, p)$. Desde que X es la suma de variable Bernoulli(p) independientes y desde que cada variable de Bernoulli tiene varianza $p(1-p)$, tenemos

$$X \sim \text{Binomial(n,p)} \rightarrow \text{Var}(X) = np(1-p).$$

Prueba de las propiedades de la varianza

Prueba de la propiedad 2: Esto se sigue de las propiedades de $\mathbb{E}(X)$ y algo de álgebra. Sea $\mu = \mathbb{E}(X)$. Entonces $\mathbb{E}(aX + b) = a\mu + b$ y

$$\begin{aligned}\mathrm{Var}(aX+b) &= \mathbb{E}((aX+b-(a\mu+b))^2) = \mathbb{E}((aX-a\mu)^2)\\ &= \mathbb{E}(a^2(X-\mu)^2) = a^2\mathbb{E}((X-\mu)^2) = a^2\mathrm{Var}(X).\end{aligned}$$

Prueba de la propiedad 3: Usamos la propiedad de $\mathbb{E}(X)$ y un poco de álgebra. Se debe recordar que μ es una constante y que $\mathbb{E}(X) = \mu$.

$$\begin{aligned}\mathbb{E}((X-\mu)^2) &= \mathbb{E}(X^2 - 2\mu X + \mu^2)\\ &= \mathbb{E}(X^2) - 2\mu\mathbb{E}(X) + \mu^2\\ &= \mathbb{E}(X^2) - 2\mu^2 + \mu^2\\ &= \mathbb{E}(X^2) - \mu^2\\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2.\end{aligned}$$