Universidad Nacional de La Plata

Facultad de Ciencias Astronómicas y Geofíscias

Estadística Aplicada

Resumen teórico

Autor:
Lorenzo Girotti

2024

Unidad 1

Probabilidad

1.1. Definiciones

Antes que nada, vamos a definir conceptos que utilizaremos a lo largo del resumen.

- **Variable aleatoria:** Es el resultado de un experimento aleatorio, que puede ser por ejemplo la acción de lanzar un dado y ver qué número queda en la cara superior. En notación, diremos que X es una variable aleatoria y la denotaremos por X.

- **Variable determinista:** Es un valor que podemos *determinar* (valga la redundancia) y usualmente asumiremos que se comportan como números reales o subconjuntos de los mismos. Las notamos como x.

- **Espacio muestral o universo:** Representa la totalidad de los posibles resultados de un experimento aleatorio, normalmente lo notaremos como E o U.

- **Eventos mutuamente excluyentes:** Un *evento* lo podemos pensar como un suceso o algo que puede ocurrir. Cuando tenemos dos o más eventos que tienen la característica de que no pueden coexistir o suceder al mismo tiempo decimos que son *mutuamente excluyentes*. Por ejemplo: sea el evento A que salga cruz; y el evento B que salga cara. A y B son eventos mutuamente excluyentes en el experimento de lanzar una moneda.

- **Probabilidad empírica:** Dado un evento que sea resultado de un experimento aleatorio que consiste en n pruebas, definimos la probabilidad de que ocurra ese evento como:

$$P(\mathrm{X}) = \frac{\#X}{n} \tag{1.1}$$

  siendo $\#X$ la cantidad de veces que ocurre el evento aleatorio X.

1.2. Reglas de probabilidad y resultados importantes

- **Probabilidad aditiva:**

$$P(A) + P(B) - P(A \cap B). \tag{1.2}$$

en caso de que sean *mutuamente excluyentes*, $P(A \cap B) = 0$

- **Probabilidad conjunta:**

$$P(A \cap B) = P(A|B)P(B) \tag{1.3}$$

si A y B son *estadísticamente independientes*, entonces $P(A|B) = P(A)$ y $P(A \cap B) = P(A)P(B)$.

- **Probabilidad total:** Dado un espacio muestral compuesto de n subconjuntos: $E = A_1 \cup \cdots \cup A_n$, la probabilidad de que ocurra un evento B dentro del espacio, es

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i), \tag{1.4}$$

considerando que es la unión de todas las intersecciones entre B y los subconjuntos A_i y asumiendo que los mismos son *mutuamente excluyentes*. Si tenemos en cuenta la 1.3, la probabilidad total queda

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i). \tag{1.5}$$

- **Teorema de Bayes:** Sean A_i como en el ítem anterior,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} \tag{1.6}$$

Unidad 2

Variable aleatoria

2.1. Función de distribución de una variable aleatoria

Si consideramos una variable aleatoria discreta, definimos a su *función de distribución* como

$$P(\mathrm{X} < x) = \sum_{i=1}^{k(x_k<x)} P(\mathrm{X} = x_i). \tag{2.1}$$

Si la variable aleatoria es continua, se define la función de densidad lineal de probabilidad:

$$f(x)dx = P(x \leq \mathrm{X} < x + dx) \tag{2.2}$$

Notar que para una función continua no tiene sentido hablar de la probabilidad de que X sea igual a un valor determinado. Sí diremos que la variable aleatoria se encuentra en un intervalo infinitesimal de posibilidades.

Podemos relacionar la *densidad de probabilidad* con la *distribución* de la variable a través de:

$$f(x) = \frac{dF(x)}{dx} \tag{2.3}$$

. Tener en cuenta que

$$f(x) \geq 0; \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

.

De esta manera, podemos calcular la probabilidad acumulada de una variable aleatoria continua a través de su función de densidad de probabilidad.

$$F(x) = \int_{-\infty}^{x} f(t)dt. \tag{2.4}$$

$$P(x_1 < \mathrm{X} < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx. \tag{2.5}$$

Notar que

$$F(-\infty) = \int_{-\infty}^{-\infty} f(x)dx = 0; \qquad F(\infty) = \int_{-\infty}^{\infty} f(x)dx = 1.$$

2.2. Operador Esperanza y parámetros de variables aleatorias

Definimos al *operador esperanza* como:

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

si X es discreta.

$$E\left[H(\mathrm{X})\right] = \int_{-\infty}^{\infty} H(x)f(x)dx$$

si X es continua.

El operador es *lineal*, es decir:

$$\begin{gathered} E\left[\mathrm{X} + \mathrm{Y}\right] = E\left[\mathrm{X}\right] + E\left[\mathrm{Y}\right] \\ E\left[c\mathrm{X}\right] = cE\left[\mathrm{X}\right], \end{gathered}$$

donde c es una constante.

2.2.1. Momentos de orden

Dado un $H(\mathrm{X}) = (\mathrm{X} - c)^l$, el valor esperado $E\left[H(\mathrm{X})\right]$ se llama *momento de orden* l de una variable con respecto a c.

Si tomamos a $c = \mu$ siendo μ la media de una variable, definimos $\mu_l = E\left[(\mathrm{X} - \mu)^l\right]$ y tenemos:

$$\begin{aligned} \mu_0 &= 1 \\ \mu_1 &= 0 \\ \mu_2 = E\left[(\mathrm{X} - \mu)^2\right] &= \sigma^2 \quad \text{Varianza} \\ \mu_3 &= E\left[(\mathrm{X} - \mu)^3\right] \quad \text{Sesgo de X} \\ \mu_4 &= E\left[(\mathrm{X} - \mu)^4\right] \quad \text{Curtosis de X} \\ &\vdots \end{aligned}$$

Nota: $E\left[\mathrm{X}\right] = \mu$

Cada *momento* define mejor el comportamiento de la variable.

2.2.2. Variable normalizada

Una variable normalizada es aquella que tiene media cero y varianza uno.

$$\mathrm{u} = \frac{\mathrm{X} - \mu}{\sigma_{\mathrm{X}}} \tag{2.6}$$

2.2.3. Desigualdad de Chebychev

$$P(|\mathrm{x} - \mu| > \mathrm{k}\sigma) \leq \frac{1}{k^2} \tag{2.7}$$

siendo $k \in \mathbb{R}$.

La desigualdad nos da una idea probabilística de nuestra variable X, conociendo solo su media y varianza (sin importar cual es su función de probabilidad o probabilidad asociada). Por ejemplo: si elegimos $k = 3$, tenemos el *criterio de los tres sigmas* para descartar *outlayers*.

2.3. Transformación de variables

Si tenemos una transformación de variable Y(X) biyectiva, *la probabilidad en intervalos equivalentes se conserva*, esto implica

$$P(x \leq \mathrm{X} < x + dx) = P(y \leq \mathrm{Y} < y + dy)$$
$$f(x)dx = g(y)dy$$

.

Entonces podemos definir la densidad de probabilidad de g como

$$g(y) = \left|\frac{dx}{dy}\right| f(x) \tag{2.8}$$

. Gráficamente podemos decir que hay una *igualdad de áreas.*

2.4. Multivariables

Todo el análisis que se realiza para una variable aleatoria, puede extenderse a una idea multivariable.

2.4.1. Función de distribución

Siendo dos variables continuas, tenemos:

Función de densidad de probabilidad conjunta

$$f(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x,y), \tag{2.9}$$

y,

$$P(a \leq \mathrm{X} < b, c \leq \mathrm{y} < d) = \int_a^b \left[\int_c^d f(x,y)dy\right] dx. \tag{2.10}$$

Función de densidad de probabilidad marginal

Se consideran todos los valores posibles para una variable, dejando una densidad de probabilidad con única dependencia de la otra.

$$g(x) = \int_{-\infty}^{\infty} f(x,y)dy \tag{2.11}$$

tenemos entonces,

$$P(a \leq \mathrm{X} < b, -\infty < y < \infty) = \int_a^b \left[\int_{-\infty}^{\infty} f(x,y)dy\right] dx = \int_a^b g(x)dx \tag{2.12}$$

Independencia de variables

Dos varaibles se dicen independientes *estadísticamente* si

$$f(x,y) = g(x)h(y) \tag{2.13}$$

donde $g(x)$ y $h(y)$ son las funciones de densidad marginales.

Probabilidad condicional

Partimos de la base de encontrar:

$$P(y \leq \mathrm{Y} < y + dy \,|\, x \leq \mathrm{X} < x + dx)$$

La función de densidad de probabilidad de Y si X:

$$f(y|x) = \frac{f(x,y)}{g(x)} \tag{2.14}$$