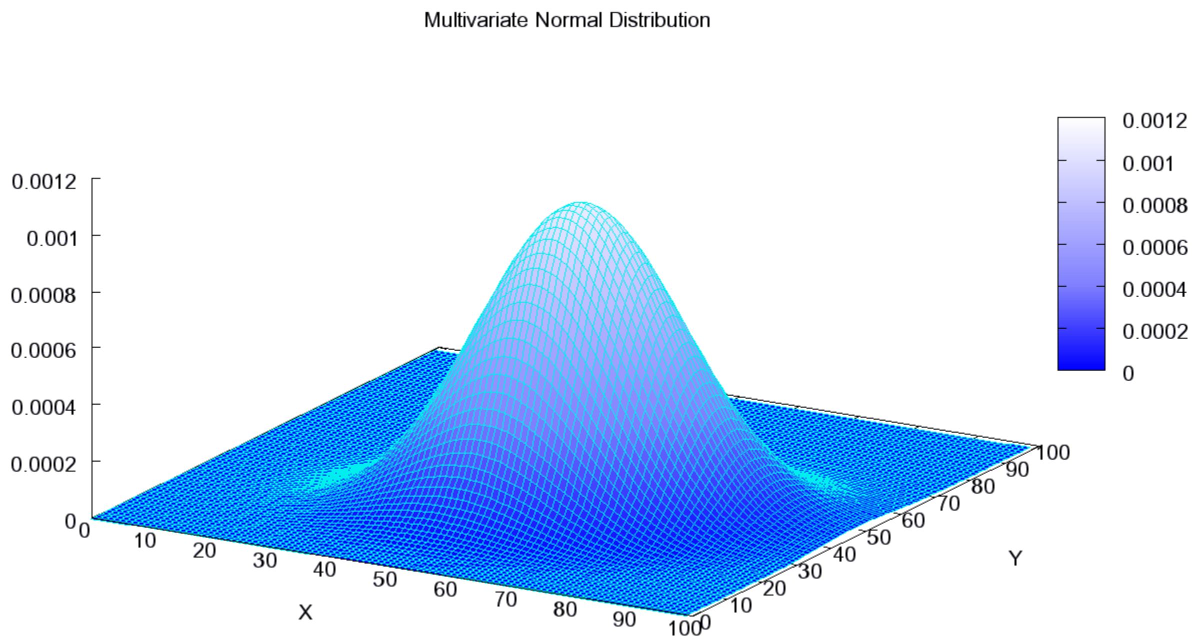
Estadística Computacional



Distribuciones Multivariadas 📈 📉

Braulio Fuentes - Diego Quezada

Temario

- Introducción.
- Vectores Aleatorios.
- Distribuciones Condicionales.
- Independencia.
- Propiedades de la Esperanza y Varianza.
- Covarianza.
- Correlación.
- Esperanza y Varianza de un Vector Aleatorio.
- Distribuciones Multivariadas
    - Distribución Multinomial
    - Distribución Normal Multivariada

Introducción.

En los capítulos anteriores estudiamos modelos de probabilidad para una sola variable aleatoria. Debido a que muchos problemas implican diversas variables aleatorias al mismo tiempo, es que vamos a estudiar modelos de probabilidad del comportamiento simultáneo de diversas variables aleatorias.

> Una compañía que comercializa latas con 3 tipos de maní, donde el peso neto de cada lata es exactamente de 1 lb. Como los tres pesos suman 1, un modelo de probabilidad conjunta de dos tipos de maní da la información necesaria sobre el peso del tercer tipo. 🥜

Vectores Aleatorios.

Un vector aleatorio es una función de Ω en $\mathbb{R}^n$, éste se puede representar de la forma

$$X = (X_1, X_2, \ldots, X_n)$$

en donde cada coordenada es una función de Ω en $\mathbb{R}$.

> Un vector aleatorio es simplemente un vector de variables aleatorias, donde cada componente del vector es una variable aleatoria. 😄

Vectores Discretos

La `fmp` de una sola variable aleatoria discreta X especifica cuánta masa de probabilidad está colocada en cada valor posible de X. Mientras que la ***función masa de probabilidad conjunta***, $p(x, y)$, de dos variables aleatorias discretas X y Y describe cuánta masa de probabilidad se coloca en cada par posible de valores (x, y).

$$p(x, y) = P(X = x \wedge Y = y)$$

debe cumplirse que $p(x, y) \geq 0$ y $\sum_x \sum_y p(x, y) = 1$.

Función Masa de Probabilidad Marginal

La función masa de probabilidad marginal de una v.a. discreta perteneciente a un *vector discreto* se obtiene sumando las probabilidades conjuntas correspondientes a todos los valores posibles de las otras v.a. Para el caso dos variables aleatorias discretas X y Y con `fmpc` $p(x,y)$:

$$p_x(x) = \sum_y p(x,y) \quad p_y(y) = \sum_x p(x,y)$$

Vectores Continuos

La probabilidad de que el valor observado de una variable aleatoria continua X esté en un intervalo se obtiene integrando la `fdp` $f(x)$ a lo largo del intervalo. Del mismo modo, la probabilidad de que el par (X, Y) de variables aleatorias continuas quede dentro un conjunto de dos dimensiones se obtiene integrando una función llamada ***función masa de probabilidad conjunta***, $f(x, y)$.

$$P[(X, Y) \in A] = \iint_A f(x, y)\, dx\, dy$$

debe satisfacer que $f(x, y) \geq 0$ y $\displaystyle\iint f(x, y)\, dx\, dy = 1$

Función de Densidad de Probabilidad Marginal

La función de densidad de probabilidad marginal de una v.a. discreta perteneciente a un *vector continuo* se obtiene integrando las probabilidades conjuntas correspondientes a todos los valores posibles de las otras v.a. Para el caso dos variables aleatorias continuas X y Y con `fdpc` $f(x, y)$:

$$f_x(x) = \int f(x, y)\, dy \quad f_y(y) = \int f(x, y)\, dx$$

Distribuciones Condicionales

Sea (X, Y) un vector aleatorio, la probabilidad condicional de Y dado que $X = x$ es:

$$\underbrace{p_{y|x} = \frac{p(x,y)}{p_x(x)}}_{\text{Caso Discreto}} \qquad \underbrace{f_{y|x} = \frac{f(x,y)}{f_x(x)}}_{\text{Caso Continuo}}$$

donde $p(x, y)$ y $f(x, y)$ son las funciones de probabilidad respectivas y $p_x(x)$ y $f_x(x)$ son las funciones de probabilidad marginal de X, las cuales deben mayor a 0.

> El caso la probabilidad condicional de X dado que $Y = y$ es análogo 💃.

Independencia.

Se dice que dos variables aleatorias X y Y son *independientes*, $X \perp Y$, si por cada par de valores x y y se cumple:

$$\underbrace{p(x,y) = p_x(x) \cdot p_y(y)}_{\text{Caso Discreto}} \qquad \underbrace{f(x,y) = f_x(x) \cdot f_y(y)}_{\text{Caso Continuo}}$$

> En caso que no se cumpla que la función de probabilidad conjunta es el producto de las probabilidades marginales se dice que las v.a. son dependientes. 🤕

Propiedades de la Esperanza y Varianza.

Sean X y Y variables aleatorias con esperanza y varianza finita y $a, b \in \mathbb{R}$.

- $E(a) = a$
- $E(aX + b) = a \cdot E(X) + b$
- $E(X + Y) = E(X) + E(Y)$

- $V(a) = 0$
- $V(aX + b) = a^2 \cdot V(X)$
- $V(X \pm Y) = V(X) + V(Y) \pm 2 \cdot Cov(X, Y)$
- Sea $X \perp Y$:
  - $E(XY) = E(X) \cdot E(Y)$
  - $V(X \pm Y) = V(X) + V(Y)$

Covarianza.

La covarianza de dos variables aleatorias X y Y, denotada por $Cov(X,Y)$, es:

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

- Si $Cov(X,Y) >> 0$, existe una fuerte relación positiva entre las v.a.
- Si $Cov(X,Y) << 0$, existe una fuerte relación negativa entre las v.a.
- Si X y Y no están fuertemente relacionadas la covarianza tendrá un valor cercano 0.

> El defecto de la covarianza es que su valor calculado depende críticamente de las unidades de medición 😢.

Propiedades de la Covarianza

Sean X y Y variables aleatorias y sea $a \in \mathbb{R}$.

- $Cov(X, Y) = E(XY) - E(X) \cdot E(Y)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, X) = Var(X)$
- $Cov(a, Y) = 0$
- $Cov(aX, Y) = a \cdot Cov(X, Y)$
- $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$
- Si $X \perp Y$, entonces $Cov(X, Y) = 0$.
- En general, $Cov(X, Y) = 0 \nRightarrow X \perp Y$.

Correlación

El coeficiente de correlación de X y Y, denotado por ρ, se define como:

$$\rho = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$$

- $-1 \leq \rho \leq 1$.
- Como regla empírica se dice que la relación es fuerte si $|\rho| \geq 0.8$, moderada si $0.5 < |\rho| < 0.8$ y débil si $|\rho| \leq 0.5$.

Interpretaciones de ρ

1. La ρ mide el grado de asociación lineal entre X y Y y sólo cuando las dos variables están perfectamente relacionadas de una manera lineal ρ será tan positivo o negativo como pueda ser.
2. Un ρ menor que 1 en valor absoluto indica sólo que la relación no es completamente lineal, sino que aún puede haber una fuerte relación no lineal.
3. Un $\rho = 0$ no implica que X y Y son independientes, sino sólo que existe una ausencia completa de relación lineal. No obstante, pueden ser altamente dependientes porque existe una fuerte relación no lineal.

Esperanza y Varianza de un Vector Aleatorio

La esperanza y varianza de una variable aleatoria pueden extenderse al caso de vectores aleatorios.

Esperanza.

Sea X el vector aleatorio $(X_1, \ldots, X_n)$. Cuando cada coordenada del vector tiene esperanza finita se define la esperanza de X como el vector numérico:

$$E(X) = \mu = (E(X_1), \ldots, E(X_n))$$

> En ocasiones suele escribirse la esperanza como un vector columna 🤓.

Varianza

Sea X el vector aleatorio $(X_1, \ldots, X_n)$. Cuando cada coordenada del vector tiene primer y segundo momento finito entonces la varianza de X se define como la matriz cuadrada

$$
\begin{aligned}
Var(X) = \Sigma &= \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \ldots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \ldots & \\ \vdots & \vdots & & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \ldots & Var(X_n) \end{pmatrix}_{n \times n} \\
&= E\left[(X - E(X))^T \cdot (X - E(X))\right]
\end{aligned}
$$

> Esta matriz también se conoce como *matriz de varianzas y covarianzas*, la cual ademas es simétrica y positiva definida 🤯.

Distribuciones Multivariadas

Distribución Multinomial

Un experimento compuesto de n ensayos independientes e idénticos, en donde cada ensayo puede dar como resultado cualquiera de r posibles resultados.

Sea p_i la probabilidad de ocurrencia del resultado i y las variables aleatorias X_i como el número de ensayos que dan el resultado i. La `fmpc` de $X_1, \ldots, X_r$ se llama **distribución multinomial**.

La función masa de probabilidad conjunta de $X_1, \ldots, X_r$ es:

$$p(x_1, \ldots, x_r) = \frac{n!}{(x_1!)(x_2!) \cdot \ldots \cdot (x_r!)} p_1^{x_1} \cdot \ldots \cdot p_r^{x_r} \quad x_i = 0, 1, 2, \ldots, n$$

donde $x_1 + \ldots + x_r = n$

- $\mu = (np_1, \ldots, np_r)$
- $[\Sigma]_{ij} = \begin{cases} np_i(1 - p_i) & \textbf{si } i = j \\ -np_i p_j & \textbf{si } i \neq j \end{cases}$

> Cuando $r = 2$ da como resultado la distribución binomial 😱.

Distribución Normal Multivariada

Se dice que el vector $X = (X_1, \ldots, X_n)$ tiene una distribución normal multivariada si su función de densidad es:

$$f(X) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right)$$

en donde μ es el vector de medias y Σ es la matriz de varianzas y covarianzas.

> Cuando el vector aleatorio tiene dos componentes se dice que tiene una *Distribución Normal Bivariada* 👌.

Propiedades Distribución Normal Bivariada

Sea (X, Y) un vector aleatorio continuo que tiene una distribución normal bivariada con correlación ρ y con esperanza y varianza finita:

- X tiene distribución marginal $N\left(\mu_x, \sigma_x^2\right)$
- Y tiene distribución marginal $N\left(\mu_y, \sigma_y^2\right)$
- $E(X, Y) = (\mu_x, \mu_y)$
- $\Sigma = \begin{pmatrix} \sigma_x^2 & \rho \cdot \sigma_x \cdot \sigma_y \\ \rho \cdot \sigma_x \cdot \sigma_y & \sigma_y^2 \end{pmatrix}$

- La probabilidad condicional de Y dado que $X = x$ tiene una distribución Normal:

$$N\left(\mu_y + \frac{\rho \cdot \sigma_y(x - \mu_x)}{\sigma_x}, (1 - \rho^2) \cdot \sigma_y^2\right)$$

- Cuando $\mu_x = \mu_y = 0$, $\sigma_x = \sigma_y = 1$ y $\rho = 0$, la distribución se llama ***normal bivariada estándar***, la cual tiene la siguiente función de densidad conjunta:

$$f(x, y) = \frac{1}{2\pi} \cdot e^{\left(\frac{-1}{2} \cdot \left(x^2 + y^2\right)\right)}$$