CAPITULO II

VARIABLES ALEATORIAS

FUNCIONES DE DENSIDAD DE PROBABILIDAD

El escenario más relevante para nosotros es cuando la realización de un proceso aleatorio se presenta en forma numérica (p.e. corresponde a una medición) : a cada elemento X corresponde una variable x (real o entera). Para x continuo, su función de densidad de probabilidad (PDF) $P(x)$ se define como :

$$\mathcal{P}(X \text{ en } [x, x + dx]) = P(x)dx ,$$

donde $P(x)$ es definida-positiva para todo de x, y satisface la condición de normalización

$$\int_{-\infty}^{+\infty} dx' P(x') = 1 .$$

Para x_i discreto, la definición es similar :

$$\mathcal{P}(X \text{ en } x_i) = p_i ,$$

$$\text{con } \sum_j p_j = 1 \text{ y } p_k \geq 0 \; \forall k .$$

Probabilidades finitas se obtienen por integración sobre un rango no-infinitesimal,

$$\mathcal{P}(a \leq X \leq b) = \int_a^b dx' P(x') .$$

Ejercicio: determinar el coeficiente de normalización de la función Gaussiana, $g(x; \mu, \sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. Y mejor aún: normalizar la Gaussiana en un intervalo truncado, con $a \leq x \leq b$.

En ocasiones es conveniente referirse a la función de densidad cumulativa (CDF) :

$$C(x) \; = \; \int_{-\infty}^{x} dx' P(x') \; ,$$

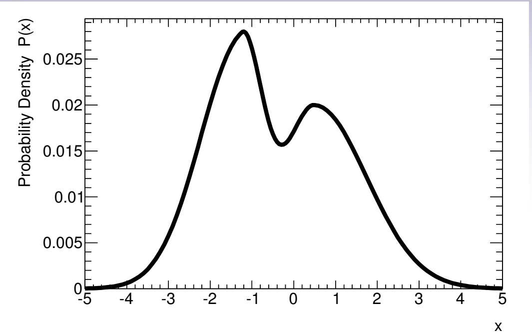
de modo que las probabilidades finitas corresponden a evaluar la CDF en los bordes del rango de interés :

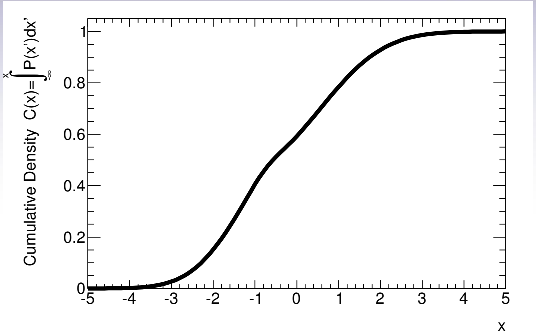
$$\mathcal{P}(a \leq X \leq b) \; = \; \int_{a}^{b} dx' P(x') \; = \; C(b) - C(a) \; .$$

Una PDF no puede ser completamente arbitraria :

- debe satisfacer la condición de normalización previamente indicada
- debe ser definida positiva
- debe ser de soporte acotado, con valores despreciables fuera de una región finita

Fuera de esas condiciones, una PDF puede ser arbitraria, p.e. exhibir uno o varios máximos locales, tener discontinuidades...
En contraste, la CDF es una función monotónicamente creciente de x.
(ver ejemplo en la lámina siguiente)





Un ejemplo arbitrario de PDF con un máximo global y un segundo máximo local, y su CDF correspondiente. Fuera del intervalo en el gráfico, el valor de la PDF es totalmente despreciable.

Ejercicio : Suponer que nuestra PDF $P(x)$ es la suma de dos PDFs Gaussianas

$$P(x;\mu_1,\sigma_1,\mu_2,\sigma_2,f) = fG(x;\mu_1,\sigma_1) + (1-f)G(x;\mu_2,\sigma_2)\ .$$

1. verificar que $P(x)$ está normalizada;
2. para los valores siguientes: $\mu_1 = -1$, $\sigma_1 = 1$, $\mu_2 = 1$, $\sigma_2 = 2$, $f = 0{,}5$. Representar gráficamente $P(x)$ y su CDF.

PDFs multidimensionales (I)

Para un evento descrito por un conjunto n-dimensional de elementos $\vec{X} = \{X_1, X_2, \ldots, X_n\}$ y su correspondiente conjunto de variables aleatorias $\vec{x} = \{x_1, x_2, \ldots, x_n\}$, tenemos su PDF multidimensional :

$$P(\vec{x})d\vec{x} \;=\; P(x_1, x_2, \ldots, x_n)dx_1 dx_2 \ldots dx_n \;.$$

PDFs de menor dimensionalidad pueden derivarse por integración de ciertas variables. Por ejemplo, para una variable específica $x = x_j$ su densidad de probability marginal unidimensional $P_X(x)$ es :

$$P_X(x)dx = dx \int_{-\infty}^{+\infty} dx_1 \ldots \int_{-\infty}^{+\infty} dx_{j-1} \int_{-\infty}^{+\infty} dx_{j+1} \ldots \int_{-\infty}^{+\infty} dy_{n-1} \;.$$

Caso bidimensional, con elementos X, Y y variables aleatorias $\vec{X} = \{x, y\}$. La probabilidad finita en un rango bidimensional rectangular es

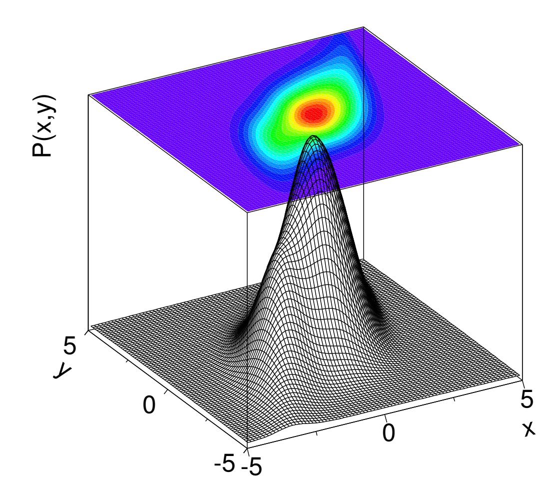
$$\mathcal{P}(a \leq X \leq b\;;\; c \leq Y \leq d) \;=\; \int_a^b dx \int_c^d dy P(x, y) \;.$$

Para un valor fijo de Y, la función de densidad condicional de X es

$$P(x \mid y) = \frac{P(x, y)}{\int dy P(x, y)} = \frac{P(x, y)}{P_Y(y)} \;.$$

De nuevo, la relación $P(x, y) = P_X(x) \cdot P_Y(y)$ solamente es válida para X, Y independientes.

Ejemplo de una función de densidad bidimensional con variables no-independientes, $P(x, y) \neq P_X(x) \cdot P_Y(y)$.



Ejercicio : Producir un gráfico que represente una PDF bidimensional razonablemente similar a la ilustrada aquí.

Modelo : descripción de un proceso aleatorio
Modelo analítico = descripción de un proceso aleatorio con funciones analíticas para las PDFs
Modelo paramétrico : sus PDFs pueden describirse completamente usando un número finito de parámetros

- este requisito no es obligatorio; las PDFs pueden también ser no-paramétricas (equivalente a suponer que se necesita un número infinito de parámetros), o pueden ser mixtas

Una implementación sencilla de una PDF paramétrica es cuando sus parámetros son argumentos analíticos de la función de densidad ; la notación

$$P\left(x, y, \ldots ; \theta_1, \theta_2, \ldots\right)$$

indica la dependencia funcional o *forma* de la PDF en términos de variables $x_1, y_2, \ldots$ y parámetros $\theta_1, \theta_2, \ldots$.

Consideremos una variable aleatoria X con PDF $P(x)$. Para una función genérica $f(x)$, su *valor de expectación* $E[f]$ es su promedio ponderado sobre el rango cubierto por x :

$$E[f] \ = \int dx P(x) f(x) \ = \ \frac{\int dx P(x) f(x)}{\int dx P(x)} \ .$$

Como describiremos más adelante, los parámetros de una PDF pueden ser estimados a partir de ciertos valores de expectación.

Por ser de uso frecuente, algunos valores de expectación tienen nombre propio.
Para PDFs unidimensionales, la *media* y la *varianza* se definen así :

$$\begin{aligned} \text{Media} \quad : \mu \; &= \quad E[x] \; = \int dx P(x) x \; , \\ \text{Varianza} \quad : \sigma^2 \; &= \quad V[x] \; = \; E[x^2] - \mu^2 \; = \; E[(x-\mu)^2] \; ; \end{aligned}$$

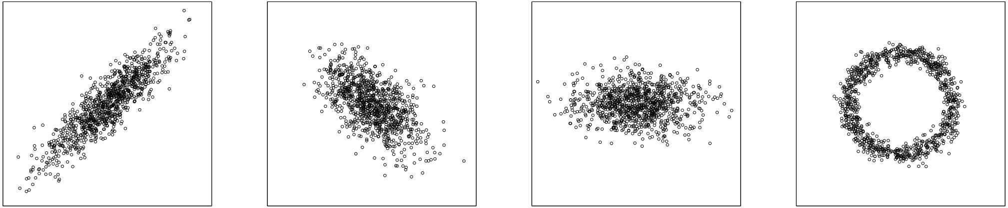
y la *desviación estándar* σ es la raíz cuadrada de la varianza.
Para PDFs multidimensionales, la matriz de *covarianza* $C_{ij} = C(x_i, x_j)$ y la matriz adimensional de correlación linear ρ_{ij} se definen así :

$$C_{ij} \; = \; E[x_i x_j] - \mu_i \mu_j \; = \; E[(x_i - \mu_i)(x_j - \mu_j)] \; , \; \rho_{ij} \; = \; \frac{C_{ij}}{\sigma_i \sigma_j} \; .$$

Los coeficientes de correlación lineal tienen valores en el rango $-1 \leq \rho_{ij} \leq 1$, e indican la tendencia dominante de densidad en el patrón $(x_j; x_j)$ pattern: se habla de correlaciones positivas y negativas (o anti-correlaciones).
Para variables aleatorias X_i, X_j independientes, es decir con $P(x_i, x_j) = P_{X_i}(x_i) P_{X_j}(x_j)$, se tiene

$$E[x_i x_j] \; = \int \int dx_i dx_j P(x_i, x_j) x_i x_j \; = \; \mu_i \mu_j \; , \; \longrightarrow \; \rho_{ij} = 0 \; .$$

(pero la conversa no es necesariamente cierta, p.e. ejemplo en la lámina siguiente)



De izquierda a derecha :

- $\rho = +0{,}9$,
- $\rho = -0{,}5$;
- $\rho = 0$, para variables independientes ,
- variables fuertemente correlacionadas con un patrón no lineal de correlación que "conspira" para arrojar una correlación lineal nula, $\rho = 0$.