# Modelado Estocástico

Prof. Fernando Grosz

fgrosz@udesa.edu.ar

Clase 1 – Comentarios Adicionales

**Algunos comentarios adicionales:**

Comentario 1:

$E(X)$, $E(X^2)$, $E(X^3)$, $E(X^4)$ se llaman el primer, segundo, tercer y cuarto momento de la variable aleatoria X, respectivamente.

Comentario 2:

$$Var(X) = E[X - E(X)]^2 = E(X^2) - \big(E(X)\big)^2$$

A la varianza también se la llama "segundo momento **central**".

Comentario 3: el coeficiente de asimetría (*skewness*) para una muestra de $n$ observaciones,

$$Skewness = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3$$

donde $\bar{x}$ y $s$ son la media y desvío estándar muestral respectivamente (es decir, hay una muestra de una cierta variable).

Ahora bien, si trabajamos con variables aleatorias, el coeficiente de asimetría de una variable aleatoria se define como:

$$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}}$$

donde $\mu_i$ denota el i-ésimo momento central, o sea, $\mu_i = E(X - E(X))^i$.

Por otro lado, la curtosis de una variable aleatoria X se define como:

$$\alpha_4 = \frac{\mu_4}{(\mu_2)^2}$$

Note que las definiciones de asimetría y curtosis vistas en la clase 1 son esencialmente las mismas, aplicadas a una muestra. Solo que ahora, con variables aleatorias continuas, hay que calcular una integral para conocer cada momento.

Comentario 4:

**La función generadora de momentos** de una variable aleatoria X se define como

$$M_X(t) = E(e^{tX})$$

en la medida que la esperanza exista en un entorno de 0, es decir, que exista un $h > 0$, tal que para todo $t$ en $-h < t < h,\ E(e^{tX})$ exista. Notemos que esta esperanza es:

Si X es discreta, $M_X(t) = \sum_x e^{tx}\, P(X = x)$

Si X es continua, $M_X(t) = \int_{-\infty}^{\infty} e^{tx}\, f_X(x)dx$

La utilidad de la función generadora de momentos es que su n-ésima derivada evaluada en $t = 0$ es el n-ésimo momento de la variable aleatoria.

Aclaración: no siempre existe la función generadora de momentos de una variable aleatoria. Pero si existe, esta función caracteriza unívocamente todos los momentos de la variable aleatoria (los infinitos momentos).

En el curso vamos a estar viendo hasta el cuarto momento.

## Comentario 5: Jarque-Bera (JB) y Q-Q plot

Recordemos que JB es un estadístico para testear la hipótesis nula de que los datos siguen una distribución normal contra la alternativa de que los datos siguen una distribución que no es normal.

Si no rechazamos la hipótesis nula, entonces podemos proceder asumiendo que los datos (por ejemplo, los retornos de un activo o portafolio) siguen una distribución normal.

Vale la pena destacar que el Jarque-Bera es un estadístico y detrás hay fundamento estadístico para afirmar o no que los datos siguen una distribución normal. Sin embargo, si rechazamos la hipótesis nula, el enfoque visual del QQ plot podría seguir siendo útil, particularmente si no estamos interesados en analizar todo el soporte de la variable aleatoria, sino, por ejemplo, una sola cola de la distribución.

# Comentario 6 sobre Mixturas de Normales

Una mixtura de normales es, como lo dice la palabra, mezclar variables aleatorias normales. Por ejemplo, si mezclamos dos normales, digamos $X_1 \sim N(0,1)$ y $X_2 \sim N(0,4)$ con pesos 25% y 75% respectivamente, obtenemos otra variable aleatoria que se llama mixtura de normales o "*normal mixture*" en inglés. Como las variables aleatorias deben integrar a 1 por definición, por eso les damos pesos a las normales que componen la mixtura que sumen 1.

Una variable aleatoria que es mixtura de normales con el mismo primer momento (pero diferente varianza) tendrá **mayor curtosis** que una variable aleatoria normal (léase, con la misma varianza que la mixtura). Esto no es nada obvio.

**Ejemplo:** mezclamos una normal estándar y otra normal N(0,4) (media cero y varianza 4) con pesos 0.5 cada una.

La pdf de la mixtura es:

$$f(x) = 0.5\varphi_1(x) + 0.5\varphi_2(x)$$

con $\varphi_1(x)$ y $\varphi_2(x)$ las pdf (densidades) de la normal estándar y la N(0,4) respectivamente.

El primer momento de la mixtura va a ser 0 y la varianza de la mixtura es 0.5*1+0.5*4=2.5. ¿Por qué?

La curtosis de la mixtura es

$$\alpha_4 = \frac{\mu_4}{(\mu_2)^2} = \frac{\mu_4}{(2.5)^2} =$$

$$= \frac{3(0.5 * 1^2 + 0.5 * 4^2)}{(2.5)^2} = 4.08$$

Vemos entonces que la curtosis de la mixtura es mayor a la curtosis de una variable aleatoria normal (que es siempre igual a 3). Si desean graficar las densidades, deberían graficar la de la mixtura contra la de la normal con la misma varianza que la mixtura para poder apreciar las colas más pesadas de la mixtura en relación a las de la normal.

**Cuidado:** mezclando dos o más normales pueden obtener básicamente lo que quieran (o sea, es posible obtener una variable aleatoria platicúrtica, o leptocúrtica; asimétrica positiva o negativa).