Vectores aleatorios.

Extensión a más de dos dimensiones

Definición: Sean $X_1,...,X_k$ variables aleatorias discretas, la **función de probabilidad conjunta** del vector aleatorio $(X_1,...,X_k)$ se define como:

$$p_{X_1,...,X_k}(x_1,...,x_k) = P(X_1 = x_1,....,X_k = x_k)$$

y, dado cualquier conjunto $A \subseteq \Re^k$,

$$P((X_1,...,X_k) \in A) = \sum_{(x_1,...,x_k)\in A} ..... \sum p_{X_1,...,X_k}(x_1,...,x_k)$$

Esta función satisface las siguientes propiedades:

- $p_{X_1,...,X_k}(x_1,...,x_k) \geq 0 \qquad \forall\ (x_1,...,x_k)$
- $\sum_{x_1} ... \sum_{x_k} p_{X_1,...,X_k}(x_1, x_2,..., x_k) = 1$

En forma similar a lo hecho para el caso bidimensional se pueden definir las **funciones de probabilidad marginal**. Por ejemplo, la función de probabilidad marginal de X_1 está dada por:

$$p_{X_1}(x_1) = \sum_{x_2} ... \sum_{x_k} p_{X_1,...,X_k}(x_1, x_2,..., x_k)$$

y la función de probabilidad marginal de (X_1, X_2) está dada por:

$$p_{X_1,X_2}(x_1, x_2) = \sum_{x_3} ... \sum_{x_k} p_{X_1,...,X_k}(x_1, x_2,..., x_k) .$$

Distribución multinomial: Es una generalización de la distribución Binomial. Supongamos que se repite *n* veces en forma independiente una experiencia, que en cada repetición hay *k* resultados posibles ($k \geq 2$), cada uno de los cuales ocurre con probabilidad p_i ($1 \leq i \leq k$) y que estas probabilidades se mantienen constantes en todas las repeticiones. Este experimento se denomina experimento multinomial. Si definimos

X_i: número de veces que ocurre el resultado i $(1 \leq i \leq k)$

la distribución conjunta de $(X_1,...,X_k)$ se denomina distribución multinomial de parámetros $n, p_1,...p_k$.

Notación: $(X_1,...,X_k) \sim M(n, p_1,...p_k)$

La correspondiente función de probabilidad conjunta está dada por

$$p_{X_1,...,X_k}(x_1,...,x_k) = \begin{cases} \dfrac{n!}{x_1!x_2!...x_k!} p_1^{x_2} p_2^{x_2} ... p_k^{x_k} & \text{si } 0 \le x_i \le n \ \forall i, \ \sum_{i=1}^{k} x_i = n \\ 0 & \text{en otro caso} \end{cases} \qquad (1)$$

En efecto, en primer lugar hay que notar que si $x_1 + x_2 + ... + x_k \neq n$, la función de probabilidad puntual es cero. Sean ahora $0 \le x_i \le n,$ tales que $x_1 + x_2 + ... + x_k = n$. Indicando por R_i ($1 \le i \le k$) cada uno de los k resultados posibles, una de las posibles configuraciones que producen x_i resultados R_i ($1 \le i \le k$), es

$$\underbrace{R_1...R_1}_{x_1}\underbrace{R_2...R_2}_{x_2}........\underbrace{R_k...R_k}_{x_k}$$

(alguno de los x_i's podría ser 0, en cuyo caso no aparecería ninguno de los correspondientes R_i).

Como hemos supuesto independencia entre las repeticiones, esa configuración tiene probabilidad $p_1^{x_1} p_2^{x_2} ....p_k^{x_k}$, pero es sólo una de las configuraciones posibles que producen x_i resultados R_i para $1 \le i \le k$.

¿Cuántas configuraciones diferentes hay?

$$\binom{n}{x_1} \cdot \binom{n-x_1}{x_2} \cdot \binom{n-x_1-x_2}{x_3} ......\binom{x_k}{x_k} = \frac{n!}{x_1!(n-x_1)!} \cdot \frac{(n-x_1)!}{x_2!(n-x_1-x_2)!} \cdot \cdot \cdot \frac{x_k!}{x_k!0!} =$$

$$= \frac{n!}{x_1!x_2!....x_k!}$$

y se obtiene la función de probabilidad dada en (1).

Observación: La distribución marginal de X_i es binomial de parámetros n y p_i para todo $1 \le i \le k$. En general, las marginales de una distribución multinomial son binomiales o multinomiales.

Ejemplo: De una urna que contiene 3 bolillas rojas, 2 negras, 4 azules y 1 blanca se extraen 12 bolillas *con reposición*. Definiendo

X_1: número de bolillas rojas
X_2: número de bolillas negras
X_3: número de bolillas azules
X_4: número de bolillas blancas

el vector (X_1, X_2, X_3, X_4) tiene distribución multinomial, es decir

$$(X_1, X_2, X_3, X_4) \sim M\left(12, \frac{3}{10}, \frac{2}{10}, \frac{4}{10}, \frac{1}{10}\right)$$

a) ¿Cuál es la probabilidad de que se obtengan 3 bolillas rojas, 5 negras, 4 azules y ninguna blanca?

$$p_{X_1,X_2,X_3,X_4}(3,5,4,0) = \frac{12!}{3!\,5!\,4!\,0!}\left(\frac{3}{10}\right)^3\left(\frac{2}{10}\right)^5\left(\frac{4}{10}\right)^4\left(\frac{1}{10}\right)^0 = 0.006$$

b) Calcular la probabilidad de obtener a lo sumo dos bolillas rojas.

Como $X_1 \sim \mathrm{Bi}\left(12, \frac{3}{10}\right)$, entonces

$$P(X_1 \le 2) = \sum_{i=0}^{2} p_{X_1}(i) = \sum_{i=0}^{2}\binom{12}{i}\left(\frac{3}{10}\right)^i\left(\frac{7}{10}\right)^{12-i} = 0.25$$

c) Calcular la probabilidad de obtener 3 bolillas rojas y 2 blancas.

Como las v.a. que nos interesan son X_1 y X_4, defino una nueva v.a. $Y = X_2 + X_3$. El vector aleatorio (X_1 , X_4 , Y) también tendrá distribución multinomial.

$$(X_1, X_4, Y) \sim M\left(12, \frac{3}{10}, \frac{1}{10}, \frac{6}{10}\right)$$

y, por lo tanto, la probabilidad pedida será

$$p_{X_1,X_4,Y}(3,2,7) = \frac{12!}{3!\,2!\,7!}\left(\frac{3}{10}\right)^3\left(\frac{1}{10}\right)^2\left(\frac{6}{10}\right)^7 = 0.06$$

Definición: El vector aleatorio $(X_1,...,X_k)$ es continuo si existe una función $f_{X_1,...,X_k}:\Re^k \rightarrow \Re^{\geq 0}$, denominada **función de densidad conjunta**, tal que

$$P\left((X_1,...,X_k)\in A\right)=\int_A\!...\!\int f_{X_1,...,X_k}(x_1,x_2,...,x_k)\,dx_1...dx_k \qquad \forall\, A\subseteq \Re^k$$

Esta función satisface las siguientes propiedades:

- $f_{X_1,...,X_k}(x_1,...,x_k)\geq 0 \qquad \forall\ (x_1,...,x_k)$
- $\int_{-\infty}^{\infty}......\int_{-\infty}^{\infty} f_{X_1,...X_k}(x_1,...x_k)\,dx_1.....dx_k=1$

En forma similar a lo hecho para el caso bidimensional se pueden definir las **funciones de densidad marginal**. Por ejemplo, la función de densidad marginal de X_1 está dada por:

$$f_{X_1}(x_1)=\int_{-\infty}^{\infty}....\int_{-\infty}^{\infty} f_{X_1,...,X_k}(x_1,x_2,...,x_k)\,dx_2....dx_k$$

y la función de densidad marginal de (X_1,X_2), está dada por:

$$f_{X_1,X_2}(x_1,x_2)=\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} f_{X_1,...,X_k}(x_1,x_2,...,x_k)\,dx_3...dx_k$$

Definición: $X_1,...,X_k$ son variables aleatorias **independientes** si y sólo si

$p_{X_1,...,X_k}(x_1,...,x_k)=p_{X_1}(x_1)...\,p_{X_k}(x_k) \quad \forall\,(x_1,...,x_k)$ en el caso discreto

$f_{X_1,...,X_k}(x_1,...,x_k)=f_{X_1}(x_1)....\,f_{X_k}(x_k) \quad \forall\,(x_1,...,x_k)$ en el caso continuo

Ejemplos: 1) En el caso de la distribución multinomial, las componentes del vector aleatorio son v.a. con distribución binomial no independientes y ésto que es intuitivo ya que su suma es constante (es igual a *n*), puede verificarse aplicando la definición.

2) Sea (X_1, X_2, X_3) un vector aleatorio con distribución uniforme en el prisma de vértices (0,0,0),(1,0,0),(0,2,0),(1,2,0),(0,0,3),(1,0,3),(0,2,3),(1,2,3), cuyo volumen es igual a 6. Entonces, su función de densidad conjunta dada por

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \begin{cases} 1/6 & \text{si } 0 \le \mathrm{x}_1 \le 1, 0 \le x_2 \le 2, 0 \le x_3 \le 3 \\ 0 & \text{en otro caso} \end{cases}$$

Es inmediato verificar que las componentes del vector son variables aleatorias independientes, ya que

$$f_{X_1}(x_1) = \begin{cases} \int_0^3 \int_0^2 \frac{1}{6} dx_2 \, dx_3 = \frac{1}{6} \cdot 6 = 1 & \text{si } x_1 \in [0,1] \\ 0 & \text{si } x_1 \notin [0,1] \end{cases}$$

$$f_{X_2}(x_2) = \begin{cases} \int_0^3 \int_0^1 \frac{1}{6} dx_1 \, dx_3 = \frac{1}{6} \cdot 3 = \frac{1}{2} & \text{si } x_2 \in [0,2] \\ 0 & \text{si } x_2 \notin [0,2] \end{cases}$$

$$f_{X_3}(x_3) = \begin{cases} \int_0^2 \int_0^1 \frac{1}{6} dx_1 \, dx_2 = \frac{1}{6} \cdot 2 = \frac{1}{3} & \text{si } x_3 \in [0,3] \\ 0 & \text{si } x_3 \notin [0,3] \end{cases}$$

entonces,

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = f_{X_1}(x_1)\, f_{X_2}(x_2)\, f_{X_3}(x_3) \qquad \forall\, (x_1,x_2,x_3)$$

Distribución de la suma de dos variables aleatorias

Sean *X* e *Y* dos v.a. de las cuáles se conoce la distribución conjunta. Estamos interesados en la distribución de la v.a. *V* = *X* + *Y*.

Consideraremos dos ejemplos, uno para el caso de un vector aleatorio discreto y otro para el caso continuo.

Ejemplos: 1) Sean $X \sim P(\lambda)$ e $Y \sim P(\mu)$, v.a. independientes, y sea V = X + Y. Claramente el recorrido de la v.a. V es el conjunto $R_V = \{0,1,2,....\}$. Sea $k \in R_V$,

$$P(X+Y=k) = \sum_{i=0}^{\infty} p_{XY}(i,k-i) = \sum_{i=0}^{k} p_X(i)\, p_Y(k-i)$$

por ser X e Y independientes. Entonces,

$$P(X+Y=k) = \sum_{i=0}^{k} \frac{e^{-\lambda}\lambda^i}{i!} \cdot \frac{e^{-\mu}\mu^{k-i}}{(k-i)!} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \lambda^i \mu^{k-i} =$$

$$= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda+\mu)^k .$$

Entonces, $V = X + Y$ tiene distribución de Poisson de parámetro $\lambda + \mu$. O sea

$$X + Y \sim P(\lambda + \mu)$$

Este resultado se extiende por inducción al caso de n v.a. : si $X_1,..., X_n$ son v.a. independientes tales que $X_i \sim P(\lambda_i)$ para i = 1,...,n, entonces $X_1 +...+ X_n \sim P(\lambda_1 + ...+\lambda_n)$.

2) Sean X e Y v.a. independientes con distribución exponencial de parámetro λ, o sea, sean $X \sim E(\lambda)$ e $Y \sim E(\lambda)$ independientes, y sea $V = X + Y$. La v.a. V toma valores en el intervalo $(0,\infty)$, por lo tanto, si $v \le 0$, $F_V(v)=0$. Sea $v > 0$,

$$F_V(v) = P(X+Y \le v) = \iint_{\{(x,y)\,/\,x+y \le v\}} f_{XY}(x,y)\,dx\,dy = \iint_{\{(x,y)\,/\,x+y \le v\}} f_X(x)\, f_Y(y)\,dx\,dy$$

pues X e Y son independientes. Entonces,

$$P(X+Y \le v) = \int_0^v \left(\int_0^{v-y} f_X(x)\, f_Y(y)\,dx \right) dy = \int_0^v \left(\int_0^{v-y} \lambda e^{-\lambda x}\, \lambda e^{-\lambda y}\,dx \right) dy =$$

$$= \int_0^v \lambda e^{-\lambda y} \left(\int_0^{v-y} \lambda e^{-\lambda x} dx \right) dy = \int_0^v \lambda e^{-\lambda y} \left(1 - e^{-\lambda (v-y)} \right) dy =$$

$$= \int_0^v \lambda e^{-\lambda y} dy - \int_0^v \lambda e^{-\lambda v} dy = 1 - e^{-\lambda v} - \lambda e^{-\lambda v} v$$

Derivando respecto de *v*, se obtiene la densidad de *V* = *X* + *Y*, que es

$$f_V(v) = \left(\lambda e^{-\lambda v} + \lambda^2 e^{-\lambda v} v - \lambda e^{-\lambda v}\right) I_{(0,\infty)}(v) = \lambda^2 e^{-\lambda v} v\, I_{(0,\infty)}(v)$$

lo que demuestra que *V* tiene distribución Gamma de parámetros (2,λ).

3) Se puede demostrar que, en general, si $X \sim \Gamma(\alpha,\lambda)$ e $Y \sim \Gamma(\beta,\lambda)$ son variables aleatorias independientes, entonces

$$X + Y \sim \Gamma(\alpha+\beta,\lambda)$$

Función generadora de momentos de la suma de v.a. independientes: Sean, en principio *X* e *Y* dos v.a. independientes, entonces la función generadora de la suma *X* + *Y* es el producto de las funciones generadoras, es decir

$$M_{X+Y}(t) = M_X(t)\, M_Y(t)$$

En efecto, si por ejemplo *X* e *Y* son dos v.a. continuas e independientes,

$$M_{X+Y}(t) = E\left(e^{t(X+Y)}\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t(x+y)} f_{XY}(x,y)\, dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{tx} e^{ty} f_X(x)\, f_Y(y)\, dx\, dy =$$

$$= \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx \int_{-\infty}^{\infty} e^{ty} f_Y(y)\, dy = E\left(e^{tX}\right) E\left(e^{tY}\right) = M_X(\mathrm{t})\, M_Y(\mathrm{t})$$

como queríamos demostrar. Para el caso discreto, se demuestra en forma similar.

Es inmediato verificar, usando inducción que, si $X_1, X_2, ..., X_n$ son v.a. independientes,

$$M_{X_1+X_2+...+X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

<u>Ejemplos:</u> 1) Demostraremos, usando funciones generadoras de momentos que si $X \sim P(\lambda)$ e $Y \sim P(\mu)$ son v.a. independientes, $X + Y \sim P(\lambda + \mu)$. En efecto,

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = e^{\lambda(e^t-1)} e^{\mu(e^t-1)} = e^{(\lambda+\mu)\,(e^t-1)}$$

y se obtiene la función generadora de momentos de una v.a. Poisson con parámetro $(\lambda + \mu)$. Recordemos que la función generadora de momentos determina la distribución de la v.a..

2) Demostraremos ahora, usando funciones generadoras de momentos que si X e Y son v.a. independientes con distribución exponencial de parámetro λ, o sea $X \sim E(\lambda)$ e $Y \sim E(\lambda)$, entonces $V = X + Y \sim \Gamma(2,\lambda)$. En efecto,

$$M_{X+Y}(t) = M_X(t) M_Y(t) = \frac{\lambda}{\lambda - t}\,\frac{\lambda}{\lambda - t} = \left(\frac{\lambda}{\lambda - t}\right)^2$$

y se obtiene la función generadora de momentos de una v.a. $\Gamma(2,\lambda)$.

Sumas y promedios de variables aleatorias

Propiedad: Sean $X_1, X_2, ..., X_n$ v.a. cualesquiera con $E(X_i) = \mu_i$ y $V(X_i) = \sigma_i^2$ y $a_1, a_2, ..., a_n$ números reales, entonces

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i\ \mu_i$$

(1)

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \sigma_i^2 + 2\sum_{i<j} a_i a_j \operatorname{cov}(X_i, X_j)$$

Dem: En primer lugar, probemos la expresión para la esperanza mediante inducción en n. Si $n = 2$, ya hemos demostrado que

$$E(a_1 X_1 + a_2 X_2) = a_1 E(X_1) + a_2 E(X_2)$$

Supongamos ahora que la expresión es cierta para $n = k$ y probémosla para $n = k + 1$.

En primer lugar,

$$E\left(\sum_{i=1}^{k+1} a_i X_i\right) = E\left(\sum_{i=1}^{k} a_i X_i + a_{k+1} X_{k+1}\right) = E(Y + a_{k+1} X_{k+1})$$

siendo $Y = \sum_{i=1}^{k} a_i X_i$. Como para $n = 2$ se cumple, se obtiene

$$E\left(\sum_{i=1}^{k+1} a_i X_i\right) = E(Y + a_{k+1} X_{k+1}) = E(Y) + a_{k+1} E(X_{k+1}) = E\left(\sum_{i=1}^{k} a_i X_i\right) + a_{k+1}\mu_{k+1}$$

y, utilizando la hipótesis inductiva

$$E\left(\sum_{i=1}^{k+1} a_i X_i\right) = \sum_{i=1}^{k} a_i \mu_i + a_{k+1}\mu_{k+1} = \sum_{i=1}^{k+1} a_i \mu_i$$

como queríamos demostrar.

Probemos ahora la expresión correspondiente a la varianza. Observemos en primer lugar que si *n* = *2*,

$$V(a_1 X_1 + a_2 X_2) = E\left(\left[(a_1 X_1 + a_2 X_2) - E(a_1 X_1 + a_2 X_2)\right]^2\right) =$$

$$= E\left(\left[(a_1 X_1 + a_2 X_2) - (a_1\mu_1 + a_2\mu_2)\right]^2\right) = E\left(\left[(a_1 X_1 - a_1\mu_1) + (a_2 X_2 - a_2\mu_2)\right]^2\right) =$$

$$= E\left(a_1(X_1 - \mu_1)\right)^2 + E\left(a_2(X_2 - \mu_2)\right)^2 + 2E\left[a_1 a_2 (X_1 - \mu_1)(X_2 - \mu_2)\right] =$$

$$= a_1^2 V(X_1) + a_2^2 V(X_2) + 2a_1 a_2 \operatorname{cov}(X_1, X_2)$$

y, por lo tanto, se satisface la expresión para *n* = 2.

En general,

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \operatorname{cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{i=1}^{n} a_i X_i\right) = E\left(\sum_{i=1}^{n} a_i X_i \cdot \sum_{j=1}^{n} a_j X_j\right) - E\left(\sum_{i=1}^{n} a_i X_i\right) E\left(\sum_{j=1}^{n} a_j X_j\right) =$$

$$= E\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j X_i X_j\right) - \left(\sum_{i=1}^{n} a_i \mu_i\right)\left(\sum_{j=1}^{n} a_j \mu_j\right) =$$

$$= \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j E(X_i X_j)\right) - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mu_i \mu_j =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \left(E(X_i X_j) - \mu_i \mu_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{cov}(X_i, X_j)$$

Teniendo en cuenta que si $i = j$, $\operatorname{cov}(X_i, X_i) = V(X_i)$ y que $\operatorname{cov}(X_i, X_j) = \operatorname{cov}(X_j, X_i)$, obtenemos el resultado que queríamos demostrar.

Corolario: Sean $X_1, X_2, ..., X_n$ v.a. **independientes** con $E(X_i) = \mu_i$ y $V(X_i) = \sigma_i^2$ y $a_1, a_2, ..., a_n$ números reales, entonces

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i \ \mu_i \qquad V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \sigma_i^2$$

Dem: Resulta inmediatamente del hecho que, por ser las v.a. independientes,

$$\operatorname{cov}(X_i, X_j) = 0 \quad \forall i \neq j .$$

Corolario: Sean $X_1, X_2, ..., X_n$ v.a. **independientes e idénticamente distribuidas** (i.i.d.) con $E(X_i) = \mu$ y $V(X_i) = \sigma^2 \ \forall \, i = 1, ..., n$ y $a_1, a_2, ..., a_n$ números reales, entonces

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \mu \sum_{i=1}^{n} a_i \qquad V\left(\sum_{i=1}^{n} a_i X_i\right) = \sigma^2 \sum_{i=1}^{n} a_i^2$$

Dem: Se verifica inmediatamente a partir del corolario anterior.

Propiedad: Sean $X_1, X_2, ..., X_n$ v.a. independientes e idénticamente distribuidas (i.i.d.) con $E(X_i) = \mu$ y $V(X_i) = \sigma^2 \ \forall \, i = 1, ..., n$, entonces

a) $E\left(\sum_{i=1}^{n} X_i\right) = n\mu \qquad V\left(\sum_{i=1}^{n} X_i\right) = n\sigma^2$

b) $E(\overline{X}) = E\left(\dfrac{\sum_{i=1}^{n} X_i}{n}\right) = \mu \qquad V(\overline{X}) = V\left(\dfrac{\sum_{i=1}^{n} X_i}{n}\right) = \dfrac{\sigma^2}{n}$

Dem: Ejercicio.