Distribución Multinomial: ejemplo de vectores aleatorios de dimensión n

En un experimento binomial, cada ensayo podría dar por resultado sólo uno de dos resultados posibles.

Ahora consideremos un experimento compuesto de n ensayos independientes e idénticos, en los que cada ensayo puede por resultado solo uno de r resultados posibles $(r \geq 2)$, cada uno con probabilidad $p_i, 1 \leq i \leq r$ de ocurrir. A este experimento lo llamamos *experimento multinomial*.

Si definimos las variables aleatorias

$$X_i : \text{número de veces que ocurre el resultado } i, (1 \leq i \leq r)$$

la distribución conjunta de $X_1, \cdots, X_r$ se llama *distribución multinomial* de parámetros $n, p_1, \cdots, p_r$

$(X_1, \cdots, X_r) \sim \mathcal{M}(n, p_1, \cdots, p_r)$

$$p_{X_1,\cdots,X_r}(x_1, \cdots, x_r) = \begin{cases} \dfrac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r} & \text{si } 0 \leq x_i \leq n, \forall i, \sum_{i=1}^{r} = 0 \\ 0, & \text{en otro caso} \end{cases}$$

Observación: La distribución marginal de X_i es binomial de parámetro n y $p_i, \forall 1 \leq i \leq r$

Ejercicio 1. En un experimento genético, se cruzan dos variedades diferentes de ciertas especies y una característica específica de la descendencia puede ocurrir solo en tres niveles *A*, *B* y *C*. De acuerdo a un modelo propuesto, las probabilidades para *A*, *B* y *C* son 1/2, 3/12 y 8/12 respectivamente. Sobre 60 descendientes, calcula:

a) La probabilidad de que 6, 18 y 36 descendientes, tengan los niveles *A*, *B* y *C* respectivamente.

b) La probabilidad de que 6 y 18 descendientes presenten los niveles *A* y *B*, sabiendo que 36 presentan el nivel *C*.

Chebyshev y Ley de los Grandes Números

Desigualdad de Chebyshev

Sean X una v.a con $E(X) = \mu$ y $V(X) = \sigma^2 < \infty$. Entonces,

$$\forall \epsilon > 0, \qquad P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

Ejercicio 2. *Encuestas*: Sea p la proporción de votantes que apoyan a un determinado candidato a un cargo. Entrevistamos a n votantes, seleccionados al azar, y registramos f_n, la proporción de ellos que apoyan al candidato. Acotar la probabilidad de que el estimador, f_n, difiera de p en más de $0,1$ para $n = 100$.

Si queremos, con una probabilidad de al menos el $95\,\%$, de que nuestra estimación sea muy precisa (especificamente que $|f_n - p| < 0,01$), ¿cuántos votantes deberían ser encuestados?

Ejercicio 3. Supongamos que se sabe que el número de artículos producidos en una fabrica durante una semana, es una variable aleatoria con media 50. Si se sabe que la varianza de la producción de una semana es 25, ¿qué se puede decir de la probabilidad de que la producción de esta semana este entre 40 y 60?

Definición: Decimos que la sucesión de variables aleatorias X_n converge a una variable aleatoria X si, $\forall \epsilon > 0$

$$\lim_{n \to \infty} P(|X_n - X| > \epsilon) = 0$$

Ley de los Grandes Números

Sean $X_1, X_2, \cdots$ variables aleatorias discretas independientes e idénticamente distribuidas (i.i.d) con $E(X_i) = \mu$ y $V(X_i) = \sigma^2 < \infty$. Entonces,

$$\lim_{n \to \infty} P(|\overline{X_n} - \mu| > \epsilon) = 0 \ \forall \epsilon > 0$$

O sea, $\overline{X_n} \xrightarrow{p} \mu$

Ejercicio 4. Dada una m.a $X_1, \cdots, X_n$,con función de densidad

$$f(x,\theta) = \theta x^{\theta-1} I_{(0,1)}^{(x)},$$

calcular la convergencia en probabilidad de $T_n = \frac{\overline{X_n}}{1-\overline{X_n}}$.

Ejercicio 5. Repetir el ejercicio anterior con $f(x,\theta) = \frac{1}{\theta} e^{-(x-2)/\theta}$ y $T_n = \overline{X_n} - 2$

Ejercicio 6. Ejercicio 6 de la práctica