TEMA 65: DISTRIBUCIONES DE PROBABILIDAD DE UNA VARIABLE ALEATORIA DISCRETA. CARACTERÍSTICAS Y TRATAMIENTO. LAS DISTRIBUCIONES BINOMIAL Y DE POISSON. APLICACIONES

TIEMPO: 61 — 62

Esquema

1) Introducción
   1.1) Caballero de la Mére + Pascal + Fermat
   1.2) J. Bernoulli + Moivre
   1.3) Gauss
   1.4) Laplace + Poisson
2 Definiciones
   2.1) Definición: variable aleatoria
   2.2) Definición: v.a. discreta
3) Variables Aleatorias discreta
   3.1) Distribución de probabilidad
      3.1.1) Definición: distribución de probabilidad
   3.2) Función de distribución
      3.2.1) Definición: distribución de distribución + propiedades
   3.3) Momentos de una variable aleatoria
      3.3.1) Definición: media + momentos no centrados
      3.3.2) Definición: momentos no centrados + σ^2
      3.3.3) Fórmulas
   3.4) Función generatriz de momentos
      3.4.1) Definición: fgm + función característica
4) Distribución Binomial
   4.1) Experimento de Bernoulli
   4.2) Definición: Binomial + propiedades
   4.3) Definición: distribución multinomial
5) Distribución de Poisson
   5.1) Definición: límite de la Binomial + fórmula + propiedades
6) Además
   6.1) Otras distribuciones discretas
   6.2) Aproximación de una empírica por una teórica

1) Introducción:

▷ En el s. XVII Antoine Gomband, más conocido como el Caballero de la Mére, propone a Pascal el siguiente problema:"al lanzar 24 veces un par de dados, ante la aparición de por lo menos un seis doble en las 24 tiradas, ¿es lo mismo apostar la misma cantidad a favor que en contra?"

▷ En relación a este problema y otros de la misma índole se inicia una relación epistolar entre Pascal y Fermat que se considera el inicio del Cálculo de Probabilidades.

▷ La historia del Cálculo de Probabilidades comienza con Jakob Bernouilli y de Moivre, quien formula y prueba por primera vez el Teorema Central del Límite. La obra póstuma de Bernouilli, quien perteneció a una extensa saga familiar suiza de matemáticos y físicos, "El arte de pronosticar" constituye el primer tratado importante acerca de probabilidades.

▷ Moivre (s. XVII - s. XVIII), matemático francés, formuló la probabilidad de sucesos independientes. Publicó varias obras donde estudiaba la probabilidad compuesta y a principios del s.XVIII aparece en su obra, por primera vez, la campana de Gauss.

▷ Karl Friedrich Gauss (s. XVIII - s. XIX) fue un matemático y físico alemán con contribuciones esenciales en el campo de la física y la astronomía. En lo concerniente a las Matemáticas, abarca un número muy amplio de ramas: desarrolla el Teorema de los Número Primos, una geometría no euclídea por primera vez, el método de los mínimos cuadrados, las leyes básicas de la probabilidad,...

▷ Entre Moivre y Gauss nos encontramos con Pierre Simón, marqués de Laplace. También mencionamos a Simeon Denis Poisson (s. XVIII - s. XIX) quien trabajó sobre los resultados de Laplace y analiza sucesos de probabilidad muy pequeña, lo que le condujo a la distribución que hoy lleva su nombre.

2) Definiciones:

▷ **<u>Definición</u>**: sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad asociado a un experimento aleatorio. Decimos que $\mathbb{X} : \Omega \longmapsto \mathbb{R}$ es una <u>variable aleatoria</u> $\longleftrightarrow$ la imagen inversa de cualquier conjunto Borel es un elemento de $\mathcal{A}$.

▷ *<u>Nota</u>*: la σ-álgebra de los conjuntos de Borel es la σ-álgebra más pequeña que contiene a todos los abiertos de $\mathbb{R}$.

▷ **<u>Proposición</u>**: son equivalentes

a) $\mathbb{X}$ es una variable aleatoria

b) La imagen inversa de cualquier abierto de $\mathbb{R}$ pertenece a $\mathcal{A}$

c) La imagen inversa de cualquier cerrado de $\mathbb{R}$ pertenece a $\mathcal{A}$

d) $\mathbb{X}^{-1}(]-\infty, x]) \in \mathcal{A}$, $\forall x \in \mathbb{R}$

▷ **<u>Definición</u>**: si suponemos que $\mathbb{X} = \mathbb{X}(a)$ toma un conjunto numerable de valores reales $\{x_n\}$ y el conjunto $A_n = \{a \in \Omega / \mathbb{X}(a) = x_n\}$ es tal que $A_n \in \mathcal{A}$, diremos que $\mathbb{X} : (\Omega, \mathcal{A}, P) \longmapsto \mathbb{R}$ es una <u>variable aleatoria discreta</u>.

3) Variables Aleatorias Discretas:

▷ <u>Distribución de Probabilidad</u>:

▷ Previamente, veamos cómo a un suceso cualquiera se le puede asignar una variable aleatoria. Sea un suceso aleatorio A, definimos la variable aleatoria indicador de A, χ_A dada por: $\chi_A = \{1$ si se presenta A al realizar el suceso; 0 si no se presenta $A\}$. Por tanto tenemos:
$P[\chi_A = 1] = P(A), P[\chi_A = 0] = P(\overline{A}) = 1 - P(A)$

Sea $\mathbb{X}$ una variable aleatoria discreta que toma valores x_n con $n = 1, 2, \ldots$. Vamos a asociarle a esta v.a. un conjunto de sucesos: sea el suceso A definido por $\emptyset \neq A_n = \{a \in \Omega \text{ tq } \mathbb{X}(a) = x_n\}$.

1) $A_n \cap A_m = \emptyset$ si $n \neq m$, pues si $a \in A_n \cap A_m \rightarrow \mathbb{X}(a) = x_n = x_m$ lo que es una contradicción. Luego los $\{A_n\}$ son incompatibles.

2) $\displaystyle\bigcup_{n=1}^{\infty} A_n = \Omega$

3) $\displaystyle P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) = P(\Omega) = 1$

Las condiciones **(1)** y **(2)** definen lo que se llama un sistema completo de sucesos y si también se verifica **(3)** tenemos una distribución de probabilidad (finita o numerable).

▷ <u>**Definición**</u>: una <u>distribución de probabilidad</u> se puede considerar como la sucesión de probabilidades que corresponden a los sucesos de un sistema completo o como la sucesión de probabilidades con que la variable aleatoria toma sus valores. $P_k = P[\mathbb{X} = x_k]$

▷ <u>**Teorema**</u>: si $\mathbb{X}$ es una v.a. discreta definida en $(\Omega, \mathcal{A}, P) \Rightarrow \mathbb{Y} = g(\mathbb{X})$ tal que $g(\mathbb{X})$ es un función real cualquiera, es otra v.a. definida en $(\Omega, \mathcal{A}, P)$.

▷ <u>Función de Distribución</u>:

▷ <u>**Definición**</u>: llamamos <u>función de distribución</u> de una v.a. discreta, y la notamos por $F(x)$, al valor $P[\mathbb{X} \leq x]$, $\forall x \in \mathbb{R}$; $F(x) = P[\mathbb{X} \leq x] = \displaystyle\sum_{x_j \leq x} P[X = x_j]$

▷ <u>**Propiedades**</u>:

1) $F(x)$ es continua por la derecha $\forall x \in \mathbb{R}$

2) $F(x)$ es creciente y no negativa

3) $\displaystyle\lim_{x \to -\infty} F(x) = 0; \lim_{x \to +\infty} F(x) = 1$

▷ *Nota*: a partir de $F(x)$ podemos conocer probabilidades como $P[a < \mathbb{X} \leq b] = F(b) - F(a)$. Además $F(x)$ tiene la ventaja de ser una función real de variable real, cosa que no suele ocurrir con la variable aleatoria $\mathbb{X}$. Además, si tenemos $F : \mathbb{R} \mapsto \mathbb{R}$ cumpliendo **(1)**, **(2)** y **(3)** $\Rightarrow F$ es la función de distribución de alguna variable aleatoria $\mathbb{X}$.

▷ Momentos de una v.a. discreta:

▷ Sea $\mathbb{X}$ una v.a. discreta definida sobre $(\Omega, \mathcal{A}, P)$ y sean x_j. $j = 1, 2 \ldots$ los valores que toma $\mathbb{X}$ con probabilidades respectivas p_j, es decir: $P[\mathbb{X} = x_j] = p_j$

▷ **Definición**: llamaremos media/esperanza matemática al número $E[\mathbb{X}] = \sum_{j=1}^{\infty} x_j \cdot p_j$ en caso de existir. La esperanza es un operador lineal, esto es: $E[a\mathbb{X} + b\mathbb{Y}] = a \cdot E[\mathbb{X}] + b \cdot E[\mathbb{Y}]$

▷ **Definición**: sea $\mathbb{X}$ una v.a. discreta definida sobre $(\Omega, \mathcal{A}, P)$ y supongamos que existe $E[\mathbb{X}^n]$, $n \in \mathbb{N}$. A dicho valor lo llamaremos momento no centrado de orden "n" de la v.a. $\mathbb{X}$.
Si $\exists E[\mathbb{X}^n] \longrightarrow \exists E[\mathbb{X}^k]$, $\forall k \leq n$. Lo notaremos por m_n.

▷ **Definición**: supongamos que existe $E[\mathbb{X}^n]$, llamamos momento centrado de orden "n" al valor $\mu_n = E[(\mathbb{X} - m_1)^n] = E[(\mathbb{X} - E[\mathbb{X}])^n]$
Si $\exists \mu n \longrightarrow \mu_k$, $\forall k \leq n$

▷ **Proposición**: de la linealidad de la esperanza se deducen las siguientes fórmulas:

- $\mu_n = m_n - \binom{n}{1} m_1 \cdot m_{n-1} + \binom{n}{2} m_1^2 \cdot m_{n-2} + \ldots + (-1)^n \binom{n}{n} m_1^n$
- $m_n = \mu_n + \binom{n}{1} m_1 \cdot \mu_{n-1} + \binom{n}{2} m_1^2 \cdot \mu_{n-2} + \ldots + \binom{n}{n} m_1^n$

▷ **Definición**: si existe $E[\mathbb{X}^2]$, μ_2 recibe el nombre de varianza de $\mathbb{X}$ y la notamos por σ^2.

▷ **Definición**: llamamos desviación típica a: $+\sqrt{\sigma^2} = \sigma$

De las fórmulas obtenemos que: $\mu_2 = \sigma^2 = E[(\mathbb{X} - E[\mathbb{X}])^2] = E[\mathbb{X}^2] - (E[\mathbb{X}])^2$

Proof. $\mu_2 = m_2 - \binom{2}{1} m_1 \cdot m_1 + \binom{2}{2} m_1^2 = m_2 - m_1^2 = E[\mathbb{X}^2] - (E[\mathbb{X}])^2$ □

▷ **Función Generatriz de Momentos**:

▷ Sea $\mathbb{X}$ una v.a. discreta definida sobre $(\Omega, \mathcal{A}, P)$.

▷ **Definición**: la función $M(t) = E[\mathrm{e}^{t\mathbb{X}}] = \sum_{k=0}^{\infty} \mathrm{e}^{t \cdot x_k}$ se llama función generatriz de momentos de la variable aleatoria $\mathbb{X}$, en caso de estar definida en un entorno de $t = 0$, $]-\epsilon, \epsilon[$

▷ Se puede demostrar que si existe la f.g.m. entonces es infinitamente derivable en $t = 0$ y, además, se obtiene que $M^{n)}(0) = E[\mathbb{X}^n] = m_n$, es decir, nos da los momentos no centrales sin más que derivar en cero.

▷ **Definición**: llamamos función característica a: $\varphi_{\mathbb{X}}(t) = E[\mathrm{e}^{it\mathbb{X}}] = M(i \cdot t)$

4) Distribución Binomial:

▷ La llamada distribución de Bernouilli considera un experimento con dos únicos resultados (éxito o fracaso) con sus respectivas probabilidades p y $q(= 1 - p)$. Si $\mathbb{X}$ es la v.a. "número de éxitos", la distribución de probabilidades es:

$$P[\mathbb{X} = 1] = p,\ P[\mathbb{X} = 0] = q = 1 - p \text{ y } \overline{x} = E[\mathbb{X}] = p,\ \sigma^2 = p \cdot q$$

▷ Por ejemplo: obtener cara al lanzar una moneda, obtener un 6 en el lanzamiento de un dado,...

Si repetimos el experimento de Bernouilli "n" veces tal que sean pruebas independientes lo que obtenemos es la llamada Distribución Binomial.
En relación al experimento aleatorio dado, consideremos el suceso A (éxito) y su complementario $\overline{A}$ (fracaso). Supongamos que $P(A) = p$, $P(\overline{A}) = 1 - p = q$. Además, la probabilidad se mantiene fija en cada observación y no depende de los resultados anteriores.

▷ **Definición**: definimos la variable binomial $\mathbb{X}$ como el número de éxitos tras "n" pruebas.

▷ **Distribución de probabilidad**: $P[\mathbb{X} = k] = \binom{n}{k} p^k \cdot q^{n-k}$

Proof. Veamos que cumple las condiciones para ser distribución de probabilidad:

a) $P[X = k] \geq 0$ por ser $p, q > 0$.

b) $\displaystyle\sum_{k=0}^{n} P[X = k] = \binom{n}{0} p^0 \cdot q^n + \ldots + \binom{n}{n} p^n \cdot q^0 = (p + q)^n = 1$

Supongamos que tenemos "k" éxitos en "n" tiradas. Una manera de obtenerlos es $A \overbrace{\cdots}^{k} A\overline{A} \overbrace{\cdots}^{n-k} \overline{A}$ y $P[A \cdots A\overline{A} \cdots \overline{A}] = p \cdots p \cdot q \cdots q = p^k \cdot q^{n-k}$. Pero cualquier reajuste de "k" éxitos y "$n - k$" fracasos nos vale, luego: $PR_n^{k,n-k} = C_k^n = \binom{n}{k}$ con lo que concluimos: $P[\mathbb{X} = k] = \binom{n}{k} p^k \cdot q^{n-k}$

□

▷ Como depende de dos parámetros ("n" y "p") la llamamos binomial de parámetros n y p y la notamos por $B(n, p)$. Como $\displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$ observamos que para el suceso seguro tenemos el binomio de Newton, de ahí el nombre de Binomial.

▷ **Función de distribución**: $F(x) = \displaystyle\sum_{k \leq x} \binom{n}{k} p^k q^{n-k}$ donde $x \in [0, n]$ y si $x < 0 \longrightarrow F(x) = 0$ y si $x > n \longrightarrow F(x) = 1$

▷ **Esperanza**: $E[\mathbb{X}] = n \cdot p$

Proof. $E[\mathbb{X}] = \displaystyle\sum_{k=0}^{n} \binom{n}{k} k \cdot p^k q^{n-k} = \Big(\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k}\Big) \cdot np = n \cdot p(p + 1 - p)^{n-1} = n \cdot p$

□

▷ **Varianza**: $\sigma^2[\mathbb{X}] = E[\mathbb{X}^2] - (E[\mathbb{X}])^2 = n \cdot p \cdot (1-p) = n \cdot p \cdot q$

Proof. $E[\mathbb{X}^2] = E[\mathbb{X}(\mathbb{X}-1)] + E[\mathbb{X}] = \displaystyle\sum_{k=0}^{n} \binom{n}{k} k(k-1) \cdot p^k q^{n-k} + np = n(n-1)p^2 + np \longrightarrow$

$\longrightarrow \sigma^2[\mathbb{X}] = n(n-1)p^2 + np - n^2p^2 = n \cdot p \cdot (1-p)$

□

▷ **Función generatriz de momentos**: $M(t) = [p \cdot \mathrm{e}^t + (1-p)]^n$

Proof. $E[\mathrm{e}^{t\mathbb{X}}] = \displaystyle\sum_{k=0}^{n} \binom{n}{k} e^{tk} \cdot p^k q^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (p\mathrm{e}^t)^k \cdot (1-p)^{n-k} = [p \cdot \mathrm{e}^t + (1-p)]^n$

□

▷ **Distribución multinomial**: es una generalización de la Binomial. Supongamos el espacio muestral $\Omega = \{A_1, \ldots, A_s\}$ que consta de "s" sucesos son probabilidades $p_1, \ldots, p_s$ y tal que $p_1 + \ldots + p_s = 1$. Supongamos que se realizan "n" pruebas. La probabilidad de que A_1 se repita k_1 veces,... ,A_s se repita k_s veces es:

$$\frac{n!}{k_1! \cdots k_s!} \cdot p_1^{k_1} \cdots p_s^{k_s}, \text{ tal que } k_1 + ... + k_s = n$$

donde $\dfrac{n!}{k_1! \cdots k_s!} = PR_n^{k_1,\ldots,k_s}$.

Luego: $P[\{A_j = k_j\}, j = 1, ..., s] = \dfrac{n!}{k_1! \cdots k_s!} \cdot p_1^{k_1} \cdots p_s^{k_s}$

▷ Por ejemplo: si lanzamos un dado de seis caras ocho veces, la probabilidad de obtener dos unos, dos cincos y el resto de las caras una vez es:

$$\frac{8!}{2!1!1!1!2!1!} \cdot \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)$$

5) Distribución de Poisson:

▷ También llamada "Ley de los sucesos raros". Se aplica en situaciones en se tiene una Binomial con $p \rightsquigarrow 0$ y $n >>$ y donde $n \cdot p$ se mantiene constante a $\lambda > 0$. La distribución de probabilidad se obtiene haciendo el límite $n \to \infty$ de la Binomial.

Proof. $P[\xi = k] = \lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} = \lim_{n\to\infty} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \{n \cdot p = \lambda > 0\} =$
$\lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k (1-\frac{\lambda}{n})^{n-k} = \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \left[\frac{1\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{\left(1-\frac{\lambda}{n}\right)^k}\right] =$
$= \{$ el primer término tiende a $\mathrm{e}^{-\lambda}$ y el segundo tiende a $1\} = \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda}$

□

▷ **Definición**: sea $\mathbb{X}$ una v.a. discreta sobre $\Omega, \mathcal{A}, P)$. Se dice que $\mathbb{X} \sim \mathcal{P}(\lambda)$ es una v.a. de Poisson de parámetro $\lambda > 0$ si ξ toma los valores $0, 1, 2, ...$ con probabilidades: $P[\xi = k] = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$

▷ Es una distribución de probabilidad:

Proof. $\sum_{k=0}^{\infty} P[\xi = k] = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda} \cdot \mathrm{e}^{\lambda} = 1$

□

La distribución de Poisson se aplica en situaciones como:

1) Número de radiaciones radioactivas
2) Número de partos triples por año
3) Número de llamadas telefónicas por minuto
4) Los primeros ejemplos aplicados fueron contando los muertos por coz en el ejército prusiano durante 20 años en 10 cuerpos de caballería.

▷ **Propiedades**:

1) Es asimétrica y cuando $\lambda \to \infty$ el coeficiente de apuntamiento se hace cero.
2) La distribución de Poisson, $\mathcal{P}(\lambda)$ se considera una buena aproximación a la $B(n, p)$ en el caso $n \cdot p < 5$ y $p < 0,1$ o $n > 100$ y $p < 0,05$. El interés de sustituir la Binomial por la Poisson radica en esta última depende de un solo parámetro, λ, mientras que la Binomial lo hace de dos: n y p.

▷ **Función de distribución**: $F(x) = \sum_{k \le x} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$

▷ **Esperanza**: $E[\mathbb{X}] = \lambda$

Proof. $E[\mathbb{X}] = \sum_{k=0}^{\infty} \left(e^{-\lambda} \frac{\lambda^k}{k!} \right) \cdot k = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda$

□

▷ **Varianza**: $\sigma^2[\mathbb{X}] = \lambda$

Proof. $E[\mathbb{X}^2] = \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} = E[\mathbb{X}(\mathbb{X}-1)] + E[\mathbb{X}] = \sum_{k=0}^{\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} + \lambda = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda =$
$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$
$\sigma^2[\mathbb{X}] = E[\mathbb{X}^2] - E[\mathbb{X}]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

□

▷ **Función generatriz de momentos**: $M(t) = e^{\lambda(e^t - 1)}$

Proof. $M(t) = E[e^{t\mathbb{X}}] = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!} = e^{-\lambda} e^{e^t \lambda} = e^{\lambda(e^t - 1)}$

□

▷ **Función característica**: $\varphi_{\mathbb{X}}(t) = e^{\lambda(e^{i \cdot t} - 1)}$

Proof. $\varphi_{\mathbb{X}}(t) = E[e^{i \cdot t\mathbb{X}}] = \sum_{k=0}^{\infty} e^{i \cdot tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{(e^{i \cdot t} \lambda)^k}{k!} = e^{-\lambda} e^{e^{i \cdot t} \lambda} = e^{\lambda(e^{i \cdot t} - 1)}$

□

6) Además:

$\triangleright$ Además de las contempladas aquí, existen otras distribuciones discretas como:

1) Distribución uniforme discreta
2) Distribución hipergeométrica
3) Distribución geométrica o de Pascal
4) Distribución binomial negativa
5) Etc.

$\triangleright$ Por último, se podría acometer la aproximación de una distribución empírica por una distribución teórica: por ejemplo, si una variable estadística satisface “a priori” las condiciones de una Binomial, ¿cuál de las binomiales será la más próxima? Para ello se necesita considerar:

a) Una medida en la distancia entre la empírica y la distribución ajustada
b) Un medio de apreciar la distancia medida

$\triangleright$ Se demuestra que, entre todas las binomiales, la que mejor se aproxima a una empírica de media $\overline{x}$, es la Binomial tal que $p = \dfrac{\overline{x}}{n}$. En el caso de la Poisson, el parámetro será la media: $\lambda = \overline{x}$