# Unidad 4 Distribución de probabilidades

4.1 Distribuciones discretas

Probabilidad y estadística
Ingeniería Industrial

### 4.1 Distribuciones discretas

Es posible combinar las ideas de probabilidad y frecuencias para obtener distribuciones de probabilidad que se parecen bastante a las distribuciones de frecuencias relativas, la diferencia más importante entre las distribuciones de probabilidad y las de frecuencia relativa, es que las distribuciones de probabilidad son probabilidades teóricas (modelo), mientras que las distribuciones de frecuencias relativas son probabilidades empíricas (muestras).

Sabemos que los espacios muestrales no son necesariamente numéricos. Cuando por ejemplo lanzamos una moneda tres veces, podemos registrar un resultado como cara-cara-sello o "ccs". En estadística, sin embargo, nos interesan los resultados numéricos, tal como el número de caras al lanzar una moneda tres veces.

**Ejemplo.**
Se tiene el experimento de lanzar una moneda tres veces. Sea la variable aleatoria:
X = número de caras obtenidas en los tres lanzamientos

a) ¿Cuál es el conjunto de resultados posibles?
b) Confeccionar la distribución de probabilidad.
c) Representar gráficamente la distribución.

**Definición:**
Una variable aleatoria es un número que depende del resultado aleatorio de un experimento. Una variable aleatoria es una regla que asigna un valor numérico (sólo uno) a cada punto en el espacio muestral de un experimento aleatorio.

Nota: normalmente se usan letras mayúsculas, y del final de abecedario, (X, Y, o Z) para denotar variables aleatorias.
Si la variable aleatoria es discreta la describimos según su distribución de probabilidades, que consiste en una lista de valores posibles de la variable y la proporción de veces que esperamos que ocurran.

La **distribución de probabilidades** de una variable aleatoria discreta X es una función (tabla o regla), denotada por p(x) o P [X=x], que asigna una probabilidad a cada valor posible de la variable aleatoria X.
Propiedades de una función de distribución, los valores de las probabilidades están entre 0 y 1 ($0\leq p(x)\leq 1$) para todo x.

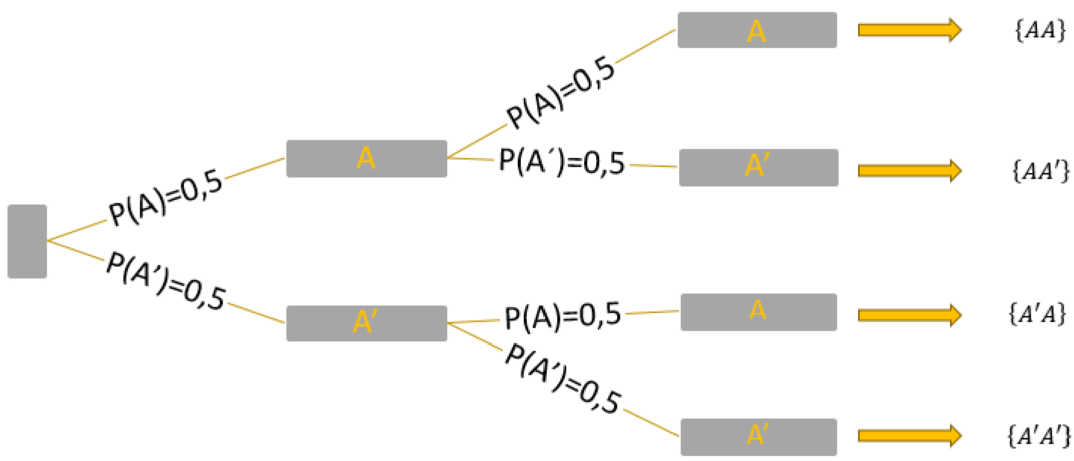
Observación: La suma de todos los valores que toma la función de probabilidad debe ser igual a 1, es decir. Notación: f.d.p. P(X=x)=p(x)

$$\sum_{i=1}^{n} P(X = x_i) = 1$$

**Ejemplo.**

Sea E: Observar el número de aciertos en el tiro al blanco, luego de 2 intentos.

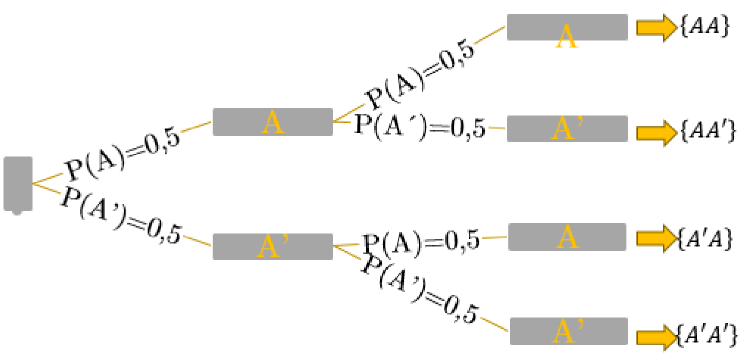
¿Cuál es el espacio muestral asociado a E?




Ω={ AA,AA',A'A,A'A'}

Sea X: Número de aciertos al blanco, luego de 2 intentos. **Rec(X)= {0,1,2}**

¿Cuáles son las probabilidades de que el número de aciertos al blanco sea de 0, 1 o 2?




| **X** | **P(X=x)** |
|---|---|
| 0 | 1/4 |
| 1 | 2/4 |
| 2 | 1/4 |

**Rec(x)**

**Función de distribución de probabilidad**

Si necesitamos calcular la probabilidad de que la variable aleatoria X tome valores menores o iguales que un cierto valor, por ejemplo, P(X≤2), es necesario acumular los distintos valores de la función de probabilidad hasta el valor deseado.

Se obtiene una nueva función acumulada, llamada función de distribución, definida por:

$$F(x) = P(X \leq x)\ \forall x \in IR$$

| $P(X \leq x)$ | X |
|---|---|
| 0 | x<0 |
| 0,25 | 0≤ x<1 |
| 0,75 | 1≤ x< 2 |
| 1 | x ≥ 2 |

**Ejemplo.**

La probabilidad del número de salidas mensuales de las personas de una faena minera viene dada por la siguiente función de probabilidad:

| **Xi: número de salidas mensuales** | **0** | **1** | **2** | **3** | **4** |
|---|---|---|---|---|---|
| **p(x)=P(X=xi)** | 0,05 | k | 0,2 | 0,2 | 0,2 |

Determine:

a) ¿Cuál es la probabilidad que una persona tenga solo un fin de semana libre?
b) ¿Cuál es la probabilidad que una persona tenga todos los fines de semana libre?
c) ¿Cuál es la probabilidad que una persona tenga al menos un fin de semana libre?
d) ¿Cuál es la probabilidad que una persona tenga a lo más un fin de semana libre?

**Esperanza de una variable aleatoria discreta**

Corresponde al valor medio (esperado) de una variable aleatoria X. Se calcula de la siguiente forma:

$$\mu = E(x) = \sum_{Rec(X)} x_i P(X = x_i)$$

La esperanza de una variable aleatoria X es un promedio ponderado, cuyas ponderaciones son las probabilidades de cada $x_i$ del Rec(X).

**Ejemplo.**

Calcule el número esperados de aciertos al blanco, luego de 2 intentos de disparos.

$$E(x) = 0 \cdot 0{,}25 + 1 \cdot 0{,}5 + 2 \cdot 0{,}25 = 1$$

**Interpretación:**

Se espera que, tras observar en un gran número de 2 intentos de disparo, que 1 acierte.

**Varianza y desviación estándar de una variable aleatoria discreta**

Varianza: Es un promedio ponderado (por las probabilidades) de los desvíos con respecto a la media al cuadrado.

Desviación estándar: Es la raíz cuadrada positiva de la varianza (la cual queda medida en las unidades originales de la variable aleatoria).

$$\sigma = \sqrt{Var(X)} = \sqrt{E(X-\mu)^2} = \sqrt{\sum_{Rec(X)} (x_i - \mu)^2 P(X = x_i)} = \sqrt{E(X^2) - E^2(X)}$$

**Ejemplo.**

Calcule la varianza y desviación estándar del número de aciertos al blanco, luego de 2 intentos de disparos.

$$\sigma^2 = \mathrm{Var}(X) = E\left[(X-\mu)^2\right] = \sum_{\mathrm{Rec}(X)} (x_i-\mu)^2 P(X=x_i) = E(X^2) - E^2(X)$$

$$E(x) = 0 \cdot 0{,}25 + 1 \cdot 0{,}5 + 2 \cdot 0{,}25 = 1$$
$$E(x^2) = 0^2 \cdot 0{,}25 + 1^2 \cdot 0{,}5 + 2^2 \cdot 0{,}25 = 1{,}5$$
$$\sigma^2 = E(x^2) - E^2(x)$$
$$\sigma^2 = 1{,}5 - 1^2$$
$$\sigma^2 = 0{,}5\ aciertos\ ^2$$

**Experimento de Bernoulli**

Un experimento de Bernoulli es aquel que sólo tiene dos resultados posibles (éxito o fracaso). Anotamos
X=1 (éxito, con probabilidad p)
X=0 (fracaso, con probabilidad q=1-p)

Ejemplos.
Lanzar una moneda y que salga cara. p=1/2.
Elegir una persona de la población y esté en DICOM. P=1/500.

Observaciones:
En experimentos con resultado dicotómico, la variable queda determinada conociendo el parámetro p.

**Distribución Binomial**

Cada ensayo de una distribución binomial termina en sólo uno de dos resultados mutuamente excluyentes, uno de los cuales se identifica como éxito y el otro como fracaso, proviene de un ensayo Bernoulli. Sin embargo, se advierte que estos términos no tienen ninguna connotación de “bueno” o “malo”.

- ✓ La probabilidad de éxito, p, para cada resultado permanece constante en un ensayo al siguiente, al igual que lo hace la de fracaso (1-p).
- ✓ La probabilidad de éxito es independiente en cada ensayo
- ✓ El experimento puede repetirse muchas veces

La probabilidad de que de n número de individuos o unidades, un x número dado se obtiene mediante la siguiente expresión:

$$P(X=x) = \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} = \binom{n}{x} p^x q^{n-x}$$

Donde:
n: tamaño de la muestra
p: probabilidad de éxito
q: probabilidad de fracaso $q = 1 - p$

La media y desviación estándar para esta distribución son:
Media o valor esperado: $\mu = n * p$ o bien $E(x) = n \cdot p$
Desviación estándar: $\sigma = \sqrt{npq}$

**Ejemplos**

Número de caras al lanzar 20 veces una moneda.

Número de aprobados si se presentan 30personas a un examen.

Número de accidentes de tráfico si han circulado 1200 automóviles.

**Ejemplo:** Se sabe que el 43,5% de los disparos realizados en un combate de fuego cruzado no dan con el objetivo.

a) Si se toma una muestra de 23 disparos. ¿Cuál es la probabilidad de que a lo más 3 disparos no den con el objetivo?

b) Si se considera una muestra de 15 disparos. ¿Cuál es la probabilidad de encontrar más de 3 pero menos de 7 disparos den con el objetivo?

c) Si se considera una muestra de 28 disparos. ¿Cuál es el número esperado de disparos que dan con el objetivo?, ¿cuál es la dispersión del número de disparos que dan con el objetivo?

**Solución:**

Se sabe que el 43,5% de los disparos realizados en un combate de fuego cruzado no dan con el objetivo.

a) Si se toma una muestra de 23 disparos. ¿Cuál es la probabilidad de que a lo más 3 disparos no den con el objetivo?

Sea X: número de disparos que no den con el objetivo.

$n = 23;\ p = 0{,}435;\ q = 1 - 0{,}435 = 0{,}565$

$X \sim B(n = 23; p = 0{,}435)$

$$\mathrm{P}(x \leq 3) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)$$

$$P(x \leq 3) = 0{,}0019363$$

$$\mathrm{P}[\mathrm{X} = 0] = \binom{23}{0} 0{,}435^0 \cdot 0{,}565^{23} = 1{,}98204 \cdot 10^{-6}$$

$$\mathrm{P}[\mathrm{X} = 1] = \binom{23}{1} 0{,}435^1 \cdot 0{,}565^{22} = 3{,}509805 \cdot 10^{-5}$$

$$\mathrm{P}[\mathrm{X} = 2] = \binom{23}{2} 0{,}435^2 \cdot 0{,}565^{21} = 2{,}972463 \cdot 10^{-4}$$

$$\mathrm{P}[\mathrm{X} = 3] = \binom{23}{3} 0{,}435^3 \cdot 0{,}565^{20} = 1{,}601973 \cdot 10^{-3}$$

Se sabe que el 43,5% de los disparos realizados en un combate de fuego cruzado no dan con el objetivo.

b) Si se considera una muestra de 15 disparos. ¿Cuál es la probabilidad de encontrar más de 3 pero menos de 7 disparos den con el objetivo?

Sea X: número de disparos que den con el objetivo.

n=15; p=0,565; q=1-0,565=0,435

$$X \sim Bin(n = 15; p = 0{,}565)$$

$$P(3 < x < 7) = P(x = 4) + P(x = 5) + P(x = 6)$$

$$P(x=4)=\binom{15}{4}\cdot 0{,}565^{4}\cdot 0{,}435^{11}=0{,}014679$$
$$P(x=5)=\binom{15}{5}\cdot 0{,}565^{5}\cdot 0{,}435^{10}=0{,}041945$$
$$P(x=6)=\binom{15}{6}\cdot 0{,}565^{6}\cdot 0{,}435^{9}=0{,}090802$$

Luego:

$$P(3<x<7)=0{,}014679+0{,}041945+0{,}090802$$
$$P(3<x<7)=0{,}147426$$

Se sabe que el 43,5% de los disparos realizados en un combate de fuego cruzado no dan con el objetivo.

c) Si se considera una muestra de 28 disparos. ¿Cuál es el número esperado de disparos que dan con el objetivo?, ¿cuál es la dispersión del número de disparos que dan con el objetivo?

Sea X: número de disparos que den con el objetivo.

Se pide el número esperado, por lo tanto, se debe calcular la esperanza y la desviación estándar.

Identificando los datos:

n=28; p=0,565; q=1-0,565=0,435

$$X\sim Bin(n=28;p=0{,}565)$$
$$E(X)=n\cdot p$$
$$E(X)=28\cdot 0{,}565$$
$$E(X)=15{,}82$$

$$\sigma=\sqrt{n\cdot p\cdot q}$$
$$\sigma=\sqrt{28\cdot 0{,}565\cdot 0{,}435}$$
$$\sigma=2{,}623$$

**Ejemplo**

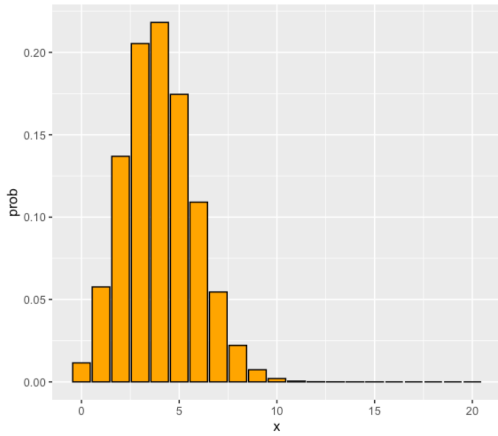
Una de las medidas de control de calidad de un amortiguador para automóvil, es probarlo en los baches de una pista, se encontró que el 20% de los amortiguadores sometidos a la prueba presentaban fuga de aceite y por lo tanto están defectuosos. Si se instalan 20 de estos amortiguadores, hallar, a) realice una representación gráfica de la distribución de probabilidades, b) la probabilidad de que más de 5 estén defectuosos. c) la probabilidad de que de 3 a 6 amortiguadores estén defectuosos.

Solución:

```
## Distribución Binomial
# a) realice una representación gráfica de la distribución de probabilidades,
x <- 0:20 # Recorrido de la variable
n <- 20 # muestra
p <- 0.2 # probabilidad de éxito
prob <- dbinom(x,n,p)
df.binom <- as.data.frame(cbind(x,prob))

library(ggplot2)
```

```r
q <- ggplot(df.binom, aes(x,prob))
q+geom_col(fill="orange", color="black")
```




```r
#b)  la probabilidad de que más de 5 estén defectuosos.
x <- 6:20
sum(dbinom(x,n,p))

## [1] 0.1957922

#c) la probabilidad de que de 3 a 6 amortiguadores estén defectuosos.
x <- 3:6
sum(dbinom(x,n,p))

## [1] 0.7072228
```

**Distribución de Poisson**

Una variable aleatoria en la medición de la frecuencia relativa de un evento sobre alguna unidad de tiempo o espacio es la distribución de Poisson. Con frecuencia se utiliza para describir el número de llegadas de clientes por hora, número de accidentes industriales en cada mes, número de fallas en una línea eléctrica, y otras.

Mide la probabilidad de un evento aleatorio sobre algún intervalo de tiempo o espacio.

Son necesarios dos supuestos para la aplicación de la distribución.

- La probabilidad de ocurrencia del evento es constante para dos intervalos de tiempo o espacio
- La ocurrencia del evento en un intervalo es independiente de la ocurrencia de otro intervalo cualquiera

La función de probabilidad de Poisson puede expresarse como:

$$P(X = x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

Esperanza: $E(x) = \mu$
Varianza=: $\sigma^2 = \mu$

**Ejemplo:** En un entrenamiento militar ocurren en promedio 3 accidentes por semana. Si estos accidentes siguen una distribución de Poisson, ¿cuál es la probabilidad de que:

a) En una semana haya 4 accidentes.
b) En un mes haya 4 accidentes.
c) En 3 días haya 4 accidentes.

En un entrenamiento militar ocurren en promedio 3 accidentes por semana. Si estos accidentes siguen una distribución de Poisson, ¿cuál es la probabilidad de que:

a) En una semana haya 4 accidentes.

Sea

X: número de accidentes en una semana.

Se conoce que $\mu = 3$

Se pide: $P(x = 4) = ?$

Luego:

$$P(x = 4) = e^{-3} \cdot \frac{3^4}{4!}$$

$$P(x = 4) = 0{,}168$$

b) En un mes haya 4 accidentes.

Sea

X: número de accidentes en un mes.

En este caso $\mu = 12$

Se pide: $P(x = 4) = ?$

Luego:

$$P(x = 4) = e^{-12} \cdot \frac{12^4}{4!}$$

$$P(x = 4) = 0{,}005308$$

c) En 3 días haya 4 accidentes.

Sea

X: número de accidentes en 3 días.

En este caso $\mu = \frac{3}{7}$

Se pide: $P(x = 4) = ?$

Luego:

$$P(x = 4) = e^{-\frac{3}{7}} \cdot \frac{\left(\frac{3}{7}\right)^4}{4!}$$

$$P(x = 4) = 0{,}0009157$$

**Ejemplo:** Supongamos que está interesado en la probabilidad de que exactamente 5 clientes lleguen durante la siguiente hora laboral. La observación simple de las últimas 80 horas ha demostrado que 780 clientes han entrado al negocio. Además, ¿Cuál es la probabilidad de que lleguen dos o menos personas? y ¿Cuál es la probabilidad de que lleguen más de tres personas?

```r
## Distribución Poisson

mu <- 780/80
# la probabilidad de que exactamente 5 clientes lleguen durante la siguiente hora laboral
x <- 5
sum(dpois(x,mu))

## [1] 0.04280265

#¿Cuál es la probabilidad de que lleguen dos o menos personas?
x <- 0:2
sum(dpois(x,mu))

## [1] 0.003397486

# ¿Cuál es la probabilidad de que lleguen más de tres personas?
x <- 0:3
sum(dpois(x,mu))

## [1] 0.01240265
```

```r
1-sum(dpois(x,mu))

## [1] 0.9875974
```