Bioestadística

Sesión 1: Modelos de probabilidad

José Aurelio Pina Romero

Ja.pina@ua.es

Bioestadística – Grado Enfermería

UA- Departamento de Enfermería

Universitat d'Alacant
Universidad de Alicante

Variable aleatoria

- El resultado de un experimento aleatorio puede ser descrito en ocasiones como una cantidad numérica.
- En estos casos aparece la noción de variable aleatoria
  - Función que asigna a cada suceso un número.
- Las variables aleatorias pueden ser discretas o continuas (como en el primer tema del curso).
- En las siguientes transparencias vamos a recordar conceptos de temas anteriores, junto con su nueva designación. Los nombres son nuevos. Los conceptos no.

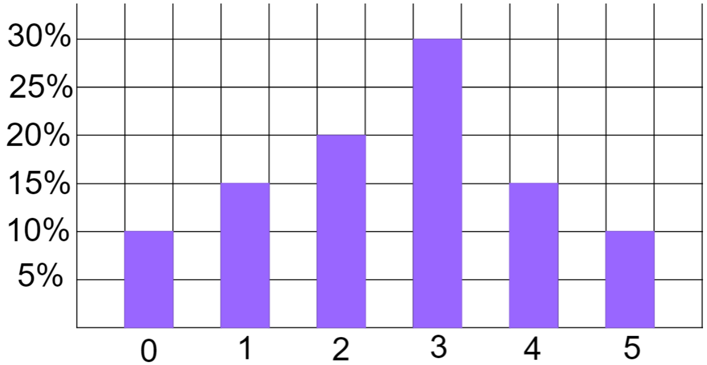
Función de probabilidad (V. Discretas)

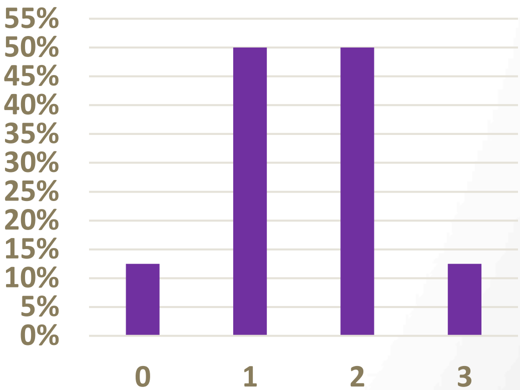
Se define como función de probabilidad de una v.a. X a una función $p(x)=P(X=x_0)$ que para cada una de los valores de la variable le asigna una probabilidad

- **Frecuencia relativa y diagrama de barras.**
- **$0 \leq P(X=x_0) \leq 1$**

Ejemplo

- **Número de caras al lanzar 3 monedas.**
- **Número de ingresos hospitalarios/día**





Función de distribución (V. Discretas)

Se define como la función de distribución de la variable X, como una función F(x) que asigna, para cada valor concreto, la probabilidad de que la variable tome un valor menor o igual a él, es decir

$$F(x_0) = P\left(X \leq x_0\right) = \sum_{x=0}^{x_0} p\left(x\right)$$

Nº ingresos	Función de probabilidad p(x)=P(X=x)	Función de distribución F(X) = P(X≤x)
0	0,10	0,10
1	0,15	0,25
2	0,20	0,45
3	0,30	0,75
4	0,15	0,90
5	0,10	1

Alguntas preguntas (V/F)

- La **función de distribución** de una variable aleatoria discreta toma valores mayores de 1 o negativos.

- La **función de distribución** de una variable aleatoria discreta toma valores entre 0 y 1.

- La **función de probabilidad** de una variable aleatoria discreta toma valores entre 0 y 1.

- La **función de probabilidad** de una variable aleatoria discreta toma valores mayores de 1.

Algunos modelos de v.a.

- **Modelos distribución discretos (Sesión 1)**

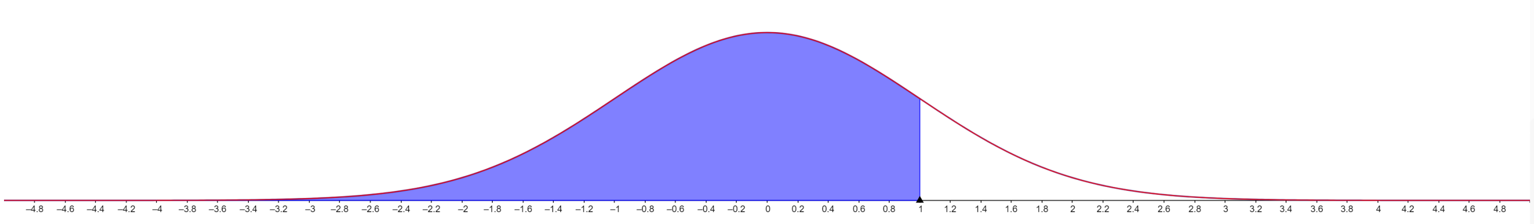
  (Cuando una función asigna probabilidad a los valores que puede tomar una v.a)

  - Bernoulli (Experimentos dicotómicos)
  - **Binomial (Contar éxitos en experimentos dicotómicos repetidos)**
  - **Poisson (sucesos raros)**

- **Modelos distribución continuos (Sesión 2)**

  (Cuando la v.a. puede tomar los infinitos valores de un intervalo)

  - **Normal**



Distribución de Bernoulli

- Tenemos un experimento de Bernoulli si al realizar un experimento sólo son posibles dos resultados:
  - X=1 (**éxito**, con probabilidad p)
  - X=0 (**fracaso**, con probabilidad q=1-p)


  - Lanzar una moneda y que salga cara.
    - p=1/2
  - Elegir una persona de la población y que esté enfermo.
    - p=1/1000 = prevalencia de la enfermedad
  - Aplicar un tratamiento a un enfermo y que éste se cure.
    - p=95%, probabilidad de que el individuo se cure

- Como se aprecia, en experimentos donde el resultado es dicotómico, la variable queda perfectamente determinada conociendo el parámetro **p**.

Ejemplo de distribución de Bernoulli.

- Se ha observado estudiando 2000 accidentes de tráfico con impacto frontal y cuyos conductores no tenían cinturón de seguridad, que 300 individuos quedaron con secuelas. Describa el experimento usando conceptos de v.a.

- Solución.
  - La noc. frecuentista de prob. nos permite aproximar la probabilidad de tener secuelas mediante 300/2000=0,15=15%
  - X="tener secuelas tras accidente sin cinturón" es variable de Bernoulli
    - X=1 tiene probabilidad $p \approx 0{,}15$
    - X=0 tiene probabilidad $q \approx 0{,}85$

Distribución binomial

Sea un **fenómeno aleatorio** que da como resultado dos sucesos. Uno con **probabilidad p** y otro con **probabilidad 1-p.**

Si se tienen **n obsevaciones** independientes del fenómeno aleatorio correspondiente a **n individuos.** La probabilidad que se observe la ocurrencia del suceso, en k de los n individuos podrá determinarse a través de una función de probabilidad binomial.

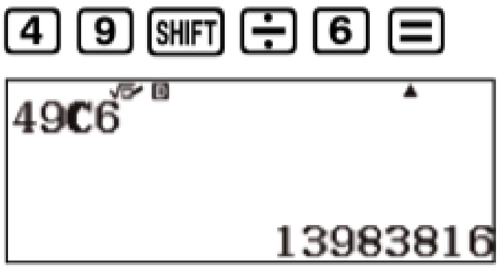
Sea X la v.a. nº de individuos en los que se observa el suceso A → **X = B(n,p)**

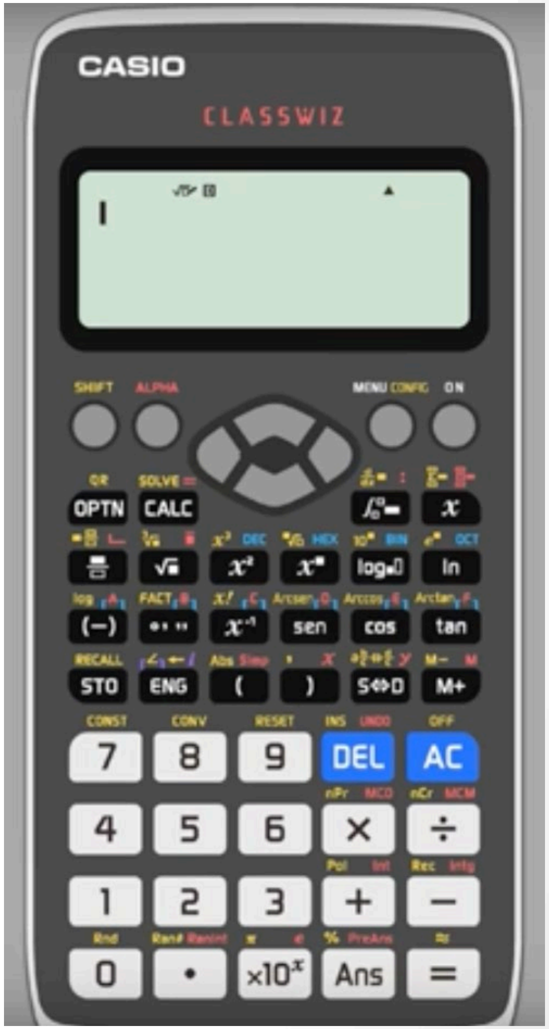
Función de probabilidad

$$P(X=k)=\binom{n}{k}p^k\,q^{n-k},\ 0\le k\le n$$

Distribución binomial

$$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$





La primitiva se trata de una combinación de 49 elementos tomados de 6 en 6.

No entran todos los elementos.
No importa el orden.
No se repiten los elementos.

Números combinatorios - Ejemplos

$$6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 =$$

$$\binom{6}{2} = \frac{6!}{(6-2)! \cdot 2!} =$$

$$\binom{6}{5} = \frac{6!}{(6-5)! \cdot 5!} =$$

$$\binom{6}{1} = \frac{6!}{(6-1)! \cdot 1!} =$$

Distribución binomial

Media: $\mu = E(X) = n \cdot p$

Varianza: $\sigma^2 = Var(X) = n \cdot p \cdot q$

https://es.wikipedia.org/wiki/Distribuci%C3%B3n_binomial

Media = número esperado

Cuadro 2.1.- Características de una variable aleatoria

Característica	Nomenclatura	Discretas	Continuas
Media o Esperanza	$\mu = E(X)$	$\sum x_i p(x_i)$	$\int_{-\infty}^{+\infty} x f(x) dx$
Varianza	$\sigma^2 = Var(X)$	$\sum (x_i - \mu)^2 p(x_i)$	$\int_{-\infty}^{+\infty} (x-\mu)^2 f(x) dx$

EJEMPLO CON CALCULADORA

Distribución binomial (n,p)

https://www.youtube.com/watch?v=Q2qjzUsdRkE

Ejercicio 1

Suponga que se conoce que el **65% de los pacientes** afectados por una patología responde positivamente al tratamiento. Si consideramos un grupo de **3 pacientes** afectados, ¿Cuál será la probabilidad que **2 de ellos** respondan de forma positiva al tratamiento?

X= número de pacientes afectados por una patología

$X \approx B(n=3, p=0,65)$

A= 2 pacientes responden + de los 3

Suceso	Paciente 1	Paciente 2	Paciente 3
A1	+	+	-
A2	+	-	+
A3	-	+	+

Ejercicio 1

Suceso	Paciente 1	Paciente 2	Paciente 3
A1	+	+	-
A2	+	-	+
A3	-	+	+

P(A1) = P(+) P(+) P(-) =
P(A2) = P(+) P(-) P(+) =
P(A3) = P(-) P(+) P(+) =

P(A) = P(A1) + P(A2) + P(A3) =

$$P(X=k)=\binom{n}{k}p^k\,q^{n-k},\ 0\le k\le n$$

Ejercicio 1

Se sabe que el **35% de los diabéticos tipo II** son tratados con insulina, si en una consulta de enfermería se dispone de **6 dosis de insulina y entran 15 sujetos.**

- i. ¿Cual es el **número esperado** de insulinodependientes?
- Ii Desviación tipica esperada de insulinodependientes.
- ii. ¿Cual es la probabilidad exacta de utilizar las seis dosis?

X= número de pacientes diabéticos tipo II tratados con insulina.

X ≈ B(n=15,p=0,35)

$$P(X=k)=\binom{n}{k}p^k\,q^{n-k},\;0\le k\le n$$

Ejercicio 2

El número de ítems a cumplimentar de un parte de declaración de una enfermedad para llevar un registro de estas es de 10. Suponiendo independiente el hecho de que un ítem no esté cumplimentado para que lo esté otro y que el **porcentaje de que un ítem no esté cumplimentado es de 0,1**

i. ¿Cuál es la probabilidad de que estén cumplimentados todos los ítems?

ii. ¿Cual es la probabilidad de que no esté cumplimentado ningún ítem?

iii. ¿Cual es la probabilidad de que 5 ítems no estén cumplimentados?

X= número de ítems a cumplimentar del parte

X = B(n=10,p=0,9)

Ejercicio 3

La probabilidad de que un estudiante del Grado de Enfermería obtenga el título es de 0,85. Si en el aula hay 80 alumnos. Calcule:

i. Probabilidad de que ninguno finalice el grado.

ii. Probabilidad de que finalicen todos.

iv. Probabilidad de que al menos 78 terminen la carrera.

iv. Halla la media y la desviación típica del número de alumnos que acaben la carrera.

X= número de alumnos del Grado de Enfermería.

X = B(n=80,p=0,85)

Ejercicio 4

Un prominente médico afirma que el 70% de los pacientes con cáncer de pulmón son fumadores empedernidos. Si su afirmación es correcta, determinar

i. La probabilidad que, de 10 pacientes recientemente ingresados en un hospital, menos de la mitad sean fumadores empedernidos.

X= número de pacientes con cáncer de pulmón

X = B(n=10,p=0,70)

Distribución de Poisson

Propuesta por **Simeón Poisson en 1837**. Útil para el estudio epidemiológico de la ocurrencia de ciertas enfermedades.

Sea λ el promedio de ocurrencias de un determinado suceso en un intervalo de tiempo o espacio. Además, supóngase que se verifican las siguientes condiciones:

1. Las ocurrencias del suceso son independientes.

2. Es posible observar un número infinito de ocurrencias en cada Intervalo. (Teóricamente).

3. La probabilidad de ocurrencia del suceso en un intervalo es proporcional a su amplitud.

La v.a. **X= número de ocurrencias** en ese intervalo de tiempo o espacio se distribuye según → **$X = P(\lambda)$**

Distribución de Poisson

Queda caracterizada por un único parámetro λ (que es a su vez su media y varianza.)

Función de probabilidad

$$P[X=k]=e^{-\lambda}\frac{\lambda^k}{k!},\ k=0,1,2,\ldots$$

$$F(X=x)=\lim_{\substack{n\to\infty \\ p\to 0 \\ np\to\lambda}}\binom{n}{x}*p^x*(1-p)^{n-x}=e^{-\lambda}*\frac{\lambda^x}{x!}$$

media = E(x)= λ

varianza= V(x)= λ

$$E(X)=\lim_{np\to\lambda} np=\lambda$$

$$\sigma^2=\lim_{np\to\lambda}(n*p*(1-p))=\lambda$$

Ejemplos

- El número de coches que pasan a través de un cierto punto en una carretera.
- El número de errores de ortografía que uno comete al escribir en una página.
- El número de llamadas telefónicas en una central telefónica por minuto (día , hora, etc)
- El número de enfermos que llegan por hora (Día , mes) a un hospital
- El número de clientes que llegan a una oficina por hora (Día , semana)
- El número de defectos por metro cuadrado de tela.
- El número de estrellas en un determinado volumen de espacio.
- El número de homicidios por día / hora en un País/región.
- El número de caso de COVID por día/semana/mes.
- El número de bacterias en 1 litro de agua.

EJEMPLO CON CALCULADORA

Distribución Poisson (λ)

https://www.youtube.com/watch?v=Q2qjzUsdRkE

Ejercicio 1

El número anual de accidentes de tránsito en el tramo de la autopista entre Alicante y Elche sigue una distribución Poisson con **una media** de **λ=10 accidentes año**.

i. Calcule la probabilidad de observar exactamente 10 accidentes en 1997.

ii. Calcule la probabilidad de observar más de 2 accidentes en 1997.

iii. Calcule la probabilidad de no observar accidentes

X= número de accidentes de tráfico

X = P(λ=10)

$$P[X=k]=e^{-\lambda}\frac{\lambda^k}{k!},\ k=0,1,2,\ldots$$

Ejercicio 2

El número de ingresos hospitalarios es de 4 ingresos-día. Para un día cualquiera,

i. Calcule la probabilidad de que se produzcan más de 5 ingresos.

ii. Calcule la probabilidad de que se produzcan entre 2 y 5 ingresos.

iii. Calcule e interprete el número esperado de ingreso en una semana y su desviación típica.

X= número de ingresos hospitalarios.

X = P(λ=4)

$$P[X=k]=e^{-\lambda}\frac{\lambda^{k}}{k!},\ k=0,1,2,\ldots$$

Ejercicio 3

El número medio de llamadas para asistencia domiciliaria a un servicio de urgencias es 7 llamadas-noche. Si sólo se pueden atender 10 llamadas.

i. Calcule la probabilidad de que se produzcan 10 llamadas en una noche.

ii. Calcule la probabilidad de que se produzcan 10 llamadas en un fin de semana (Sábado y Domingo)

iii. ¿Cuál es el valor esperado de llamadas en una noche? ¿y la desviación estándar?

X= número de llamadas hospitalarias.

X = P(λ=7)