Introducción

El presente documento tratará de realizar un análisis en profundidad de los contenidos de la materia conocida como “La materia que te desaprueba 5 veces antes de resolver el parcial”, desde ya puede tener errores, pero el mayor error es no leerlo.

1.1 - Espacios muestrales y eventos.

Simbolizamos con e a un experimento aleatorio.
Un experimento aleatorio tiene las siguientes características:
1-Se lo puede repetir bajo las mismas condiciones tantas veces como se desee.
2- No se puede predecir con exactitud el resultado de dicho experimento, pero se puede decir cuáles son los posibles resultados del mismo.
3- A medida que el experimento se repite, los resultados individuales parecen ocurrir en forma caprichosa. Pero si el experimento se repite un gran número de veces, y registramos la proporción de veces que ocurre un determinado resultado, veremos que esa proporción tiende a estabilizarse en un valor determinado a medida que aumenta el número de veces que se repite el experimento.

Se llama evento o suceso a todo subconjunto del espacio muestral. Por ejemplo,
a) En el experimento dado en el ejemplo a), un evento de S sería A = {2,4,6} pues A Ì S . Podemos expresar al evento A con palabras de la siguiente manera A: “sale un número par”.También B = {1,2,3} es un evento al que podemos expresar verbalmente como B: “sale un número menor o igual que 3”

Es útil conocer algunas leyes de Morgan:

1- Leyes de idempotencia
a) $A \cup A = A$ b) $A \cap A = A$

2- Leyes asociativas
a) $A \cup (B \cup C) = (A \cup B) \cup C$ b) $A \cap (B \cap C) = (A \cap B) \cap C$

3- Leyes conmutativas
a) $A \cup B = B \cup A$ b) $A \cap B = B \cap A$

4- Leyes distributivas
a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

5- Leyes de identidad
a) $A \cup \varnothing = A$ b) $A \cap \varnothing = \varnothing$
c) $A \cup U = U$ d) $A \cap U = A$

6- Leyes de complemento
a) $A \cup A^C = U$ b) $A \cap A^C = \varnothing$
c) $(A^C)^C = A$ d) $U^C = \varnothing$, $\varnothing^C = U$

7- Leyes de De Morgan
a) $(A \cup B)^C = A^C \cap B^C$ b) $(A \cap B)^C = A^C \cup B^C$

Definición axiomática de probabilidad.

Sea ε un experimento aleatorio y S un espacio muestral asociado con ε. Con cada evento A asociamos un número real llamado probabilidad de A, que anotamos P(A), el cual satisface las siguientes propiedades básicas o axiomas

1- $0 \le P(A) \le 1$
2- $P(S) = 1$
3- Si A y B son eventos mutuamente excluyentes entonces $P(A \cup B) = P(A) + P(B)$
4- Si $A_1, A_2, \ldots, A_n, A_{n+1}, \ldots$ es una secuencia de eventos tales que

$A_i \cap A_j = \varnothing$ si $i \neq j$, entonces $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

La elección de estas propiedades está motivada por las características correspondientes de la frecuencia relativa.

Propiedades de la probabilidad

1- $P(\varnothing) = 0$

Dem.) Siendo A un evento cualquiera podemos escribir $A \cup \varnothing = A$
Además A y $\varnothing$ son mutuamente excluyentes, por lo tanto por axioma 3

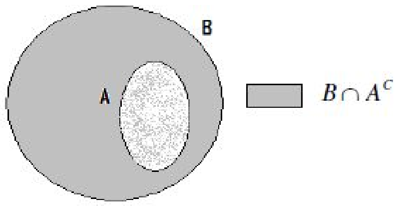
$$P(A) = P(A \cup \varnothing) = P(A) + P(\varnothing)$$

O sea que

$$P(A) = P(A) + P(\varnothing) \Rightarrow P(\varnothing) = 0$$

2- Si A^C es el evento complementario de A, entonces $P(A^C) = 1 - P(A)$

3- Si $A \subset B$ entonces $P(A) \le P(B)$



Dem.) Sean A y B dos eventos tales que $A \subset B$.
De la figura vemos que $B = A \cup (B \cap A^C)$
Y además A y $B \cap A^C$ son mutuamente excluyentes. Entonces

$P(B) = P(A) + P(B \cap A^C)$
Y como por axioma 1 tenemos que $P(B \cap A^C) \ge 0$, entonces $P(B) \ge P(A)$

4- Si $A \subset B$, entonces $P(B - A) = P(B) - P(A)$

Dem.) Siguiendo los pasos de la demostración anterior llegamos a

$P(B) = P(A) + P(B \cap A^C)$, lo que implica que $P(B \cap A^C) = P(B) - P(A)$

Y como $B \cap A^C = B - A$, entonces $P(B - A) = P(B) - P(A)$.

1.2 - El espacio muestral finito

Sea ε un experimento aleatorio y S un espacio muestral asociado a ε . Supongamos que S es un conjunto finito al que anotamos S $\{a_1, a_2, \ldots, a_n\}$, es decir consideramos que tiene n elementos. Podemos escribir a S como unión de eventos elementales de la siguiente forma. Si $A_i = \{a_i\}$ entonces $S = A_1 \cup A_2 \cup A_3 \ldots \cup A_n$. Además sabemos que P(S) = 1, por lo tanto tenemos el resultado:

$$P(A_1) + P(A_2) + \ldots P(A_n) = 1$$

Es decir: la suma de las probabilidades de los eventos elementales es igual a 1

1.3 - Espacios muestrales infinitos

Sea S un espacio muestral infinito numerable es decir S $\{a_1, a_2, \ldots, a_n\}$. Como en el caso finito, a cada a_i asignamos un número $p_i = P(\{a_i\})$ tal que

$$a)\ p_i \geq 0 \qquad b)\ \sum_{i=1}^{\infty} p_i = 1$$

La probabilidad P(A) de un evento A es entonces la suma de las probabilidades de los eventos elementales que lo componen.
Los únicos espacios muestrales infinitos no numerables que consideraremos son aquellos de medida geométrica finita m(S) tales como longitud, área o volumen, y en los cuales un punto se selecciona al azar. La probabilidad de un evento A, esto es aquella en la que el punto seleccionado pertenece a A, es entonces la relación entre m(A) a m(S) , es decir

$$P(A) = \frac{\text{longitud de } A}{\text{longitud de } S} \quad \text{ó} \quad P(A) = \frac{\text{área de } A}{\text{área de } S} \quad \text{ó} \quad P(A) = \frac{\text{volumen de } A}{\text{volumen de } S}$$

2 - Probabilidad condicional

Probabilidad condicional es la probabilidad de que ocurra un evento A, sabiendo que también sucede otro evento B. La probabilidad condicional se escribe P(A|B) o P(A/B), y se lee «la probabilidad de A dado B». No tiene por qué haber una relación causal o temporal entre A y B. A puede preceder en el tiempo a B, sucederlo o pueden ocurrir simultáneamente. A puede causar B, viceversa o pueden no tener relación causal. Las relaciones causales o temporales son nociones que no pertenecen al ámbito de la probabilidad. Pueden desempeñar un papel o no, dependiendo de la interpretación que se le dé a los eventos.

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \qquad \text{si } P(A) \neq 0$$

Análogamente

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \qquad \text{si } P(B) \neq 0$$

Teorema de la multiplicación

Si A y B son dos eventos entonces

$$P(B/A)=\frac{P(A\cap B)}{P(A)} \qquad \text{si } P(A)\neq 0$$

Por lo tanto:

$$P(A\cap B)=P(B/A)\,P(A) \qquad (6)$$

Análogamente de $P(A/B)=\dfrac{P(A\cap B)}{P(B)}$ si $P(B)\neq 0$, se deduce

$$P(A\cap B)=P(A/B)\,P(B) \qquad (7)$$

(6) y (7) se conocen como teorema de la multiplicación.

2.2 - Teorema de la probabilidad total

Sean A_1 , A_2 ,..., A_n eventos de un espacio muestral S que cumplen

a) $A_1\cup A_2\cup\ldots\cup A_n=S$

b) $A_i\cap A_j=\varnothing$ si $i\neq j$

c) $P(A_i)>0$ $\forall i=1,2,\ldots,n$ Se dice que $A_1, A_2,\ldots, A_n$ forman una ***partición de S*** Entonces para cualquier evento B de S

$$P(B)=P(B/A_1)P(A_1)+P(B/A_2)P(A_2)+\ldots+P(B/A_n)P(A_n)$$

El Teorema de la probabilidad total nos permite calcular la probabilidad de un suceso a partir de probabilidades condicionadas:

Ejemplo: supongamos que si llueve la probabilidad de que ocurra un accidentes es x% y si hace buen tiempo dicha probabilidad es y%. Este teorema nos permite deducir cuál es la probabilidad de que ocurra un accidente si conocemos la probabilidad de que llueva y la probabilidad de que haga buen tiempo.

La fórmula para calcular esta probabilidad es:

$$P(B)=\Sigma\,(A_i) * P(B/A_i) \quad \text{(donde "i" toma valores entre 1 y n)}$$

Es decir, la probabilidad de que ocurra el suceso B (en nuestro ejemplo, que ocurra un accidente) es igual a la suma de multiplicar cada una de las probabilidades condicionadas de este suceso con los diferentes sucesos A (probabilidad de un accidente cuando llueve y cuando hace buen tiempo) por la probabilidad de cada suceso A.

Teorema de Bayes

El teorema de Bayes parte de una situación en la que es posible conocer las probabilidades de que ocurran una serie de sucesos Ai. A esta se añade un suceso B cuya ocurrencia proporciona cierta información, porque las probabilidades de ocurrencia de B son distintas según el suceso Ai que haya ocurrido. Conociendo que ha ocurrido el suceso B, la fórmula del teorema de Bayes nos indica cómo modifica esta información las probabilidades de los sucesos Ai.

Sean A_1, A_2 ,..., A_n eventos de un espacio muestral S que cumplen

a) $A_1 \cup A_2 \cup \ldots \cup A_n = S$

b) $A_i \cap A_j = \varnothing \quad$ si $i \neq j$

c) $P(A_i) > 0 \quad \forall i = 1,2,\ldots,n$

Entonces para cualquier evento B de S tal que $P(B) > 0$

$$P(A_k / B) = \frac{P(B / A_k)P(A_k)}{\sum_{i=1}^{n} P(B / A_i)P(A_i)} \qquad k = 1,\ldots,n$$

Mi apreciación sobre el teorema de bayes: Si analizamos la fórmula descrita anteriormente podremos darnos cuenta que la lógica del teorema sigue siendo la misma que en muchos casos de probabilidad:

Casos favorables / Casos totales

Para poder mejorar el entendimiento de la fórmula es importante volver a repasar los conceptos del teorema de la multiplicación y el teorema de probabilidad total.

2.3 - Independencia

Puede suceder que P(B | A) y P(B) sean iguales, en ese caso A y B son eventos independientes, saber que A ocurrió no afecta la probabilidad de ocurrencia de B.

Entonces, dos eventos A y B son independientes si P(B | A) = P(B) o P(A ∩ B)=P(A).P(B)

3 - Variables Aleatorias

3.1- Generalidades

En muchas situaciones experimentales se quiere asignar un número real a cada uno de los elementos del espacio muestral. Al describir el espacio muestral de un experimento un resultado individual no tiene que ser necesariamente un número, por ejemplo, al tirar una moneda y tomar como espacio muestral S = {c, s} , o al tirar un dado dos veces tomamos como espacio muestral a S = {1,2,3,4,5,6}×{1,2,3,4,5,6} , aquí S es un conjunto de pares ordenados.

Definición: Sea ε un experimento aleatorio y S un espacio muestral asociado a él. Una variable aleatoria es una función que asigna a cada elemento de S un número real.

Notación: se anota a una variable aleatoria con letras mayúsculas X, Y, Z, W,…
Ejemplos:
1- Se tira una moneda tres veces
Sea X la v.a. X: “número de caras obtenidas luego de los tres tiros”
Si tomamos como espacio muestral
S = {(c,c,c);(c,c, s);(c, s,c);(s,c,c); (c, s, s); (s, s,c); (s,c, s); (s, s, s)}

De la misma forma que los eventos
A = {(c,c,c)} y B = {X = 3} son equivalentes
A = {(c, c, s); (c, s, c); (s, c, c)} y B = {X = 2} son equivalentes
En general:

siendo $A \subset S$ y $B \subset R_X$, A y B *son equivalentes* si $A = \{s \in S;\ X(s) \in B\}$

Las variables aleatorias se clasifican según su rango.
Sea X es una v.a. con rango Rx . Si Rx es un conjunto finito o infinito numerable entonces se dice que X es una **v.a. discreta**. Si Rx es un conjunto infinito no numerable entonces X es una **v.a. continua.**

3.2 - Variables aleatorias discretas

Se denomina variable aleatoria discreta aquella que sólo puede tomar un número finito de valores dentro de un intervalo.
Anotamos su rango como Rx ={x_1 ,x_2 ,...,x_n } si el rango es un conjunto finito de n elementos, y anotamos Rx ={x_1 ,x_2 ,... }si el rango es un conjunto infinito numerable.
A cada x_i se le asigna un número p(x_i)= P(X=Xi). Estos números deben satisfacer las siguientes condiciones:

a) $p(x_i) \geq 0$ para todo i
b) $\sum_i p(x_i) = 1$

La función $p(x)$ que antes se definió, se llama *función de probabilidad o de frecuencia de la v.a. X.*
El conjunto de pares $(x_i, p(x_i))$ $i = 1,2,...$ es la *distribución de probabilidad de X.*

Por ejemplo
1-Se tira una moneda normal tres veces, sea la v.a. X: “número de caras obtenidas”
Entonces Rx ={0 ,1,2,3}
Para hallar la distribución de probabilidad de X supongamos que la probabilidad de salir cara es ½ entonces

P(X=0)=P({(s,s,s)}) = ⅛ #Probabilidad de que salgan cero caras

P(X=1)=P({(c,s,s),(s,c,s),(s,s,c)}) = ⅜ #Probabilidad de que salgan una cara

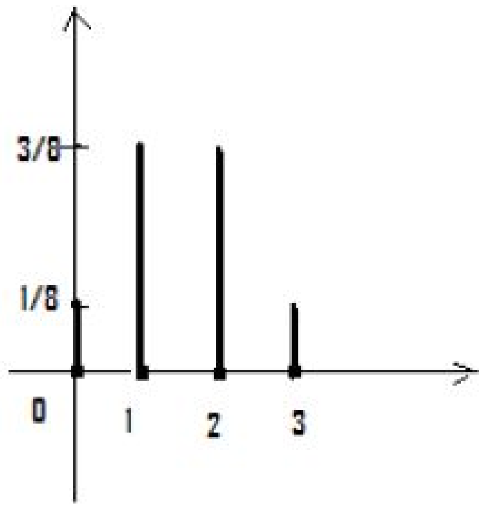
P(X=2)=P({(c,c,s),(s,c,c),(c,s,c)}) = ⅜ #Probabilidad de que salgan dos caras

P(X=3)=P({(c,c,c)}) = ⅛ #Probabilidad de que salgan tres caras

Se puede presentar la distribución de probabilidad de X en una tabla de la siguiente forma

x	0	1	2	3
p(x)	1/8	3/8	3/8	1/8

Un gráfico de la distribución de probabilidad de X sería:



Función de distribución acumulada

La función de distribución acumulada de una variable aleatoria X, evaluada en x, es la probabilidad de que X tome un valor menor o igual que x.

F(x) = P(X <= x) − ∞ < x < ∞

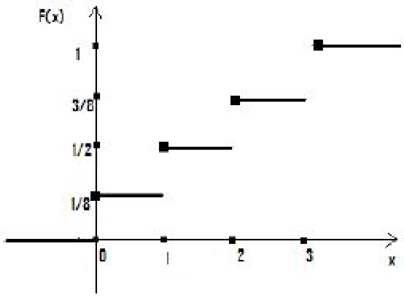
En el caso de ser X una v.a. discreta

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p(x_i) \qquad -\infty < x < \infty$$

Volviendo al ejemplo 1 anterior, la F.d.a. de X es

$$F(x)=\begin{cases} 0 & \text{si } x<0 \\ \frac{1}{8} & \text{si } 0\le x<1 \\ \frac{1}{8}+\frac{3}{8} & \text{si } 1\le x<2 \\ \frac{1}{8}+\frac{3}{8}+\frac{3}{8} & \text{si } 2\le x<3 \\ 1 & \text{si } x>3 \end{cases} \Rightarrow F(x)=\begin{cases} 0 & \text{si } x<0 \\ \frac{1}{8} & \text{si } 0\le x<1 \\ \frac{1}{2} & \text{si } 1\le x<2 \\ \frac{7}{8} & \text{si } 2\le x<3 \\ 1 & \text{si } x>3 \end{cases}$$

La gráfica de la F.d.a. de X es



<u>Observación</u>: la F.d.a. de X es una función escalonada, los puntos de "salto" coinciden con los puntos del rango de X, y la magnitud del salto en x_i es igual a $P(X = x_i)$
En general si X es una v.a. discreta cualquiera, su F.d.a. será una función escalonada.

Además si $x_1, x_2, \ldots, x_n$ son los valores del rango de X ordenados de menor a mayor entonces

$$P(X=x_1)=F(x_1)$$
$$P(X=x_i)=F(x_i)-F(x_{i-1}) \qquad i=2,\ldots,n$$

Es decir, se puede obtener la función de distribución de X a partir de su F.d.a.

3.3Esperanza de una variable aleatoria discreta

Sea X una v.a. discreta con rango Rx. La esperanza , valor medio o valor esperado de X , lo anotamos E(X) , y se define como

$$E(X)=\sum_{x_i\in R_X} x_i P(X=x_i)$$

La sumatoria se hace sobre todos los posibles valores de X
Otra notación usual es μ_X o μ

Es decir la esperanza es igual a la suma de la probabilidad de cada posible suceso aleatorio multiplicado por el valor de dicho suceso.

Ejemplos:
1- Sea la v.a. X: “número que queda en la cara de arriba al tirar un dado normal”
Rx ={1,2,3,4,5,6}

$$\text{Entonces } E(X)=\sum_{x=1}^{6} xP(X=x)=$$
$$=1P(X=1)+2P(X=2)+3P(X=3)+4P(X=4)+5P(X=5)+6P(X=6)=$$
$$=1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6}=\frac{7}{2}=3.5$$

Esperanza de una función

A veces importa hallar la esperanza de una función de X y no de X misma. Veamos un ejemplo. Un instructor de escritura técnica ha solicitado que cierto reporte sea entregado a la semana siguiente, agregando la restricción de que cualquier reporte que sobrepase las cuatro páginas será rechazado.
Sea X: “número de páginas del reporte de cierto estudiante seleccionado al azar”.
Supongamos que X tenga la siguiente distribución de probabilidad

x	1	2	3	4
p(x)	0.01	0.19	0.35	0.45

Suponga que el instructor tarda √X minutos calificando un trabajo que consiste en X páginas. Claramente √X **es otra variable aleatoria**.¿Cuál será su esperanza?, es decir ¿a qué es igual E(√X)?
Para calcular la esperanza de una v.a. se necesita conocer su función de distribución de probabilidad, por lo tanto habría que hallar previamente la distribución de probabilidad de la v.a. Y=√X. Está claro que si el rango de X es Rx = {1,2,3,4} entonces el rango de Y será Ry= {1, 2,3,4}.
Además
P(Y = √1) = P(X = 1) = 0.01
P(Y = √2) = P(X = 2) = 0.19
P(Y = √3) = P(X = 3) = 0.35
P(Y = √4) = P(X = 4) = 0.45
Por lo tanto
E(Y) = √1× P(Y = √1) + √2 × P(Y = √2) + √3 × P(Y = √3) + √4× P(Y = √4) = √1× P(X = 1) + √2 × P(X = 2) + √3 × P(X = 3) + √4 × P(X = 4) = 1.78491

O sea

$$E(Y)=\sum_x \sqrt{x}P(X=x)$$

Lo visto en este ejemplo se puede generalizar en el siguiente

Teorema: Si X es una v.a. discreta con rango R_X y distribución de probabilidad $p(x)$, entonces la esperanza de cualquier función $h(X)$ es igual a

$$E(h(X))=\sum_x h(x)p(x)$$

Propiedades de la esperanza

Supongamos un ejemplo donde Y es una función lineal de X , es decir
Y=aX+b con a y b siendo números reales. En ese caso vale entonces la siguiente propiedad

E(aX + b) = aE(X) + b

Ejemplo:

En el ejemplo anterior donde Y = 800X − 900
Directamente calculamos:
E(Y) = 800E(X) − 900
Y
E(X) = 0× 0.1+1× 0.2 + 2 × 0.3 + 3× 0.4 = 2
En consecuencia
E(Y) = 800E(X) − 900 = 800× 2 − 900 = 700

Varianza de una variable aleatoria

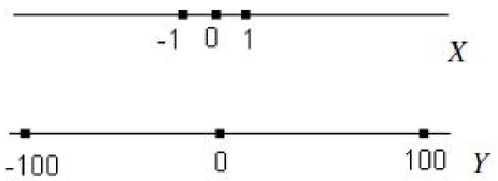
La Varianza de una variable aleatoria es una medida de dispersión definida como la esperanza del cuadrado de la desviación de dicha variable respecto a su media. Tiene como valor mínimo 0.

La esperanza de una v.a. mide dónde está centrada la distribución de probabilidad. Pero supongamos el siguiente ejemplo
Sean X e Y dos variables aleatorias con distribuciones dadas por

x	-1	1
$p(x)$	0.5	0.5

y	-100	100
$p(y)$	0.5	0.5

Es fácil verificar que $E(X) = E(Y) = 0$, pero los valores que toma la v.a. Y están más "alejados" de su esperanza que los valores de X.



Se busca una medida que refleje este hecho, se define entonces la varianza de una v.a.

Sea X una v.a. discreta con rango R_X, función de distribución de probabilidad *p(x)* y esperanza $E(X) = \mu$,

Entonces la ***varianza de X***, que anotamos $V(X)$, σ^2 o σ_X^2 es

$$V(X) = E\left[(X-\mu)^2\right] = \sum_{x \in R_X} (x-\mu)^2\, p(x)$$

La ***desviación estándar de X*** es $\sigma_X = \sqrt{V(X)}$

Otra forma de escribir la fórmula anterior:

$$V(X) = E(X^2) - \mu^2$$

Observaciones:

1- La varianza de una v.a. nunca es negativa

2- La cantidad $h(X) = (X-\mu)^2$ es el cuadrado de la desviación de X desde su media, y la varianza de X es la esperanza de la desviación al cuadrado. Si la mayor parte de la distribución de probabilidad está cerca de μ, entonces σ^2 será relativamente pequeña. Si hay valores de la variable alejados de μ que tengan alta probabilidad, entonces σ^2 será grande.

3- σ^2 está expresado en las unidades de medida de X al cuadrado, mientras que σ está expresada en las mismas unidades de medida que X.

Propiedades de la varianza

Las propiedades de la varianza de una v.a. **son consecuencia de las propiedades de la esperanza de una v.a.** Si X es una v.a. discreta con rango Rx y distribución de probabilidad p(x), entonces la varianza de cualquier función h(X) es igual a

$$V(h(X)) = \sum_{x \in R_X} \left(h(x) - E(h(X))\right)^2 p(x)$$

Si $h(X)$ es una función lineal, entonces

$$V(aX + b) = a^2 V(X) \quad \text{y} \quad \sigma_{aX+b} = \sqrt{V(aX + b)} = |a|\sigma_X$$

Observaciones:
1- $V(aX) = a^2 V(X)$
2- $V(X + b) = V(X)$

Ejemplo:
En un ejemplo anterior donde X: "número de computadoras vendidas" y Y: "utilidad obtenida", la $V(Y)$ sería $V(Y) = 800^2 V(X)$
Necesitamos calcular $V(X)$
$V(X) = E(X^2) - \mu^2$

Sabemos ya que $\mu = E(X) = 2$

Calculamos $E(X^2) = 0^2 \times 0.1 + 1^2 \times 0.2 + 2^2 \times 0.3 + 3^2 \times 0.4 = 5$

En consecuencia

$V(Y) = 800^2 V(X) = 800^2 \left(E(X^2) - \mu^2\right) = 800^2 \left(5 - 2^2\right) = 800$

3.4 Variables aleatorias discretas importantes

Distribución binomial

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos P(A) = p .
Supongamos un experimento aleatorio 0 e que cumple los siguientes requisitos:
1- se realizan n repeticiones independientes de ε , donde n se fija de antemano.
2- las repeticiones son idénticas, y en cada repetición de ε observamos si ocurre A o no ocurre A (cuando A ocurre se dice que se obtuvo un "éxito", caso contrario se obtuvo un "fracaso")
3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p
Se dice entonces que 0 ε es un experimento binomial.

Para que una variable aleatoria se considere que sigue una distribución binomial, tiene que cumplir las siguientes propiedades:

- En cada ensayo, experimento o prueba solo son posibles dos resultados (éxito o fracaso).
- La probabilidad del éxito ha de ser constante. Esta se representa mediante la letra p.

- La probabilidad de fracaso ha de ser también constante. Esta se representa mediante la letra q = 1-p. Es importante fijarse que mediante esa ecuación, sabiendo p o sabiendo q, podemos obtener la que nos falte.
- El resultado obtenido en cada experimento es independiente del anterior.
- Los sucesos son mutuamente excluyentes, es decir, no pueden ocurrir los 2 al mismo tiempo.
- Los sucesos son colectivamente exhaustivos, es decir, al menos uno de los 2 ha de ocurrir.
- La variable aleatoria que sigue una distribución binomial se suele representar como X~(n,p). n representa el número de ensayos o experimentos y p la probabilidad de éxito.

La fórmula para calcular la distribución normal es:

$$P_{(x)} = \binom{n}{x} p^x q^{n-x}$$

Donde:
n = número de ensayos/experimentos
x = número de éxitos
p = probabilidad de éxito
q = probabilidad de fracaso (1-p)

Observación: Es interesante que en la mayoría de los experimentos binomiales, interesa el número total de éxitos, más que saber exactamente cuáles repeticiones produjeron los éxitos

Esperanza y varianza

$$\boxed{\text{Sea } X \sim B(n,p), \text{ entonces } E(X) = np \text{ y } V(X) = np(1-p)}$$

Desviación típica

$$\sigma = \sqrt{n \cdot p \cdot q}$$

Distribución geométrica

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos P(A) = p .
Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:
1- se realizan repeticiones independientes de ε, hasta que ocurre A por primera vez inclusive.

2- las repeticiones son idénticas, y en cada repetición de ε observamos si ocurre A o no ocurre A(cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p

Sea la v.a. X: “número de repeticiones de ε hasta que ocurre A por primera vez inclusive”
Veamos cuál es la distribución de probabilidad de X
El rango es el conjunto $R_X = \{1,2,3,....\} = N$
Además si anotamos A_i :"ocurre A en la $i-$ésima repetición de ε" , entonces

$$P(X=k) = P(A_1^C \cap A_2^C \cap \ldots \cap A_{k-1}^C \cap A_k) = P(A_1^C)P)(A_2^C)\ldots P(A_{k-1}^C)P(A_k) = (1-p)^{k-1}p$$

por independencia

En consecuencia

$$\boxed{P(X=k) = (1-p)^{k-1}p \qquad k = 1,2,....}$$

Esperanza y varianza

Sea $X \sim G(p)$ entonces $E(X) = \dfrac{1}{p}$ y $V(X) = \dfrac{1-p}{p^2}$

Distribución binomial negativa

Esta distribución puede considerarse como una extensión o ampliación de la distribución geométrica .**Es un modelo adecuado para tratar aquellos procesos en los que se repite un determinado ensayo o prueba hasta conseguir un número determinado de resultados favorables por vez primera** .Es por tanto de gran utilidad para aquellos muestreos que procedan de esta manera. Si el número de resultados favorables buscados fuera 1 estaríamos en el caso de la distribución geométrica . Está implicada también la existencia de una dicotomía de resultados posibles en cada prueba y la independencia de cada prueba o ensayo, o la reposición de los individuos muestreados.

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos P(A) = p .
Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:
1- se realizan repeticiones independientes de ε, hasta que ocurre A por r-ésima vez inclusive.
2- las repeticiones son idénticas, y en cada repetición de e observamos si ocurre A o no ocurre A (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p

$$P(X=k)=\binom{k-1}{r-1}p^{r}(1-p)^{k-r} \qquad k=r,r+1,r+2,\ldots.$$

Notación: $X \sim BN(r,p)$

Ejemplo:

En una prueba de fuerza de soldadura, 80% de las pruebas da como resultado ruptura de soldadura, mientras que otro 20% da ruptura de la viga. Sea la v.a. X:"número de pruebas hasta la tercera ruptura de la viga inclusive". ¿Cuál es la distribución de X?. Determinar la $P(X=8)$

Solución:

Tenemos que $X \sim BN(3,\ 0.2)$

Por lo tanto $P(X=8)=\binom{8-1}{3-1}p^{3}(1-p)^{8-3}=\binom{7}{2}0.2^{3}\times 0.8^{5}=0.05505$

Esperanza y varianza

Si $X \sim BN(r,p)$ entonces $E(X)=\dfrac{r}{p}$ y $V(X)=\dfrac{r(1-p)}{p^{2}}$

Distribución hipergeométrica

Supongamos que tenemos una población o conjunto de N objetos o individuos (es decir tenemos una población finita). Clasificamos a los objetos de la población en dos categorías. Hay M objetos de una categoría y N-M de la otra categoría. Se suele decir que tenemos M "éxitos" y N-M "fracasos". Se extraen al azar y sin reemplazo n objetos de dicha población. Es decir se extrae una muestra de n objetos de la población, de manera tal que es igualmente probable que se seleccione cada subconjunto de tamaño n.
Consideramos la v.a. X: "número de éxitos en la muestra extraída"
Se dice que X tiene una distribución hipergeométrica con parámetros n, M y N

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \qquad \max(0,n-N+M)\le k\le \min(n,M)$$

Observación:

En el ejemplo anterior si el cargamento hubiese tenido 400 elementos se podría haber considerado en la parte a) a X con distribución binomial con parámetros $n = 5$ y $p = \frac{5}{400}$

En general, si el tamaño de la población N y el número de éxitos M crecen pero de manera tal que $\frac{M}{N} \to p$ y n es chico comparado con N, se puede verificar que

$$\lim_{N \to \infty} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{donde } \frac{M}{N} = p$$

Esperanza y varianza

Si $X \sim H(n, M, N)$ entonces $E(X) = \frac{nM}{N}$ y $V(X) = n\frac{M}{N}\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$

Distribución de Poisson

La distribución de Poisson es una distribución de probabilidad discreta que expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo

Una v.a. X con rango $R_X = \{0,1,2,....\}$ se dice tener ***distribución de Poisson con parámetro*** λ, si para algún $\lambda > 0$

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad k = 0,1,2,.....$$

Esperanza y varianza

Si $X \sim P(\lambda)$ entonces $E(X) = \lambda$ y $V(X) = \lambda$

Aplicaciones de la distribución de Poisson

La v.a. Poisson tiene un gran rango de aplicaciones, una de ellas es la aproximación para una v.a. binomial con parámetros n y p cuando n es grande y p es pequeño de manera tal que $np \to \lambda$, específicamente, sea $X \sim B(n, p)$ y sea $\lambda = np$, entonces

$$P(X=k)=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}=\frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}=\frac{n(n-1)....(n-k+1)}{n^k}\frac{\lambda^k}{k!}\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

Para n grande y p chico

$$\left(1-\frac{\lambda}{n}\right)^n \approx e^{-\lambda} \quad ; \quad \frac{n(n-1)....(n-k+1)}{n^k}\approx 1 \quad ; \quad \left(1-\frac{\lambda}{n}\right)^k \approx 1$$

Entonces, para n grande y p chico

$$P(X=k)\approx e^{-\lambda}\frac{\lambda^k}{k!}$$

Es decir cuando n es grande, p chico y np es "moderado" entonces la v.a. binomial con parámetros n y p tiene una distribución que se ***aproxima*** a la de una Poisson con parámetro $\lambda = np$

Tabla de Poisson

λ x	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,5	2	2,5	3	3,5	4	λ x
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679	0,2231	0,1353	0,0821	0,0498	0,0302	0,0183	0
1	0,9953	0,9825	0,9631	0,9384	0,9098	0,8781	0,8442	0,8088	0,7725	0,7358	0,5578	0,4060	0,2873	0,1991	0,1359	0,0916	1
2	0,9998	0,9989	0,9964	0,9921	0,9856	0,9769	0,9659	0,9526	0,9371	0,9197	0,8088	0,6767	0,5438	0,4232	0,3208	0,2381	2
3	1,0000	0,9999	0,9997	0,9992	0,9982	0,9966	0,9942	0,9909	0,9865	0,9810	0,9344	0,8571	0,7576	0,6472	0,5366	0,4335	3
4		1,0000	1,0000	0,9999	0,9998	0,9996	0,9992	0,9986	0,9977	0,9963	0,9814	0,9473	0,8912	0,8153	0,7254	0,6288	4
5				1,0000	1,0000	1,0000	0,9999	0,9998	0,9997	0,9994	0,9955	0,9834	0,9580	0,9161	0,8576	0,7851	5
6							1,0000	1,0000	1,0000	0,9999	0,9991	0,9955	0,9858	0,9665	0,9347	0,8893	6
7										1,0000	0,9998	0,9989	0,9958	0,9881	0,9733	0,9489	7
8											1,0000	0,9998	0,9989	0,9962	0,9901	0,9786	8
9												1,0000	0,9997	0,9989	0,9967	0,9919	9
10													0,9999	0,9997	0,9990	0,9972	10
11													1,0000	0,9999	0,9997	0,9991	11
12														1,0000	0,9999	0,9997	12
13															1,0000	0,9999	13
14																1,0000	14

λ x	4,5	5	5,5	6	6,5	7	7,5	8	8,5	9	9,5	10	10,5	11	11,5	12	λ x
0	0,0111	0,0067	0,0041	0,0025	0,0015	0,0009	0,0006	0,0003	0,0002	0,0001	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0
1	0,0611	0,0404	0,0266	0,0174	0,0113	0,0073	0,0047	0,0030	0,0019	0,0012	0,0008	0,0005	0,0003	0,0002	0,0001	0,0001	1
2	0,1736	0,1247	0,0884	0,0620	0,0430	0,0296	0,0203	0,0138	0,0093	0,0062	0,0042	0,0028	0,0018	0,0012	0,0008	0,0005	2
3	0,3423	0,2650	0,2017	0,1512	0,1118	0,0818	0,0591	0,0424	0,0301	0,0212	0,0149	0,0103	0,0071	0,0049	0,0034	0,0023	3
4	0,5321	0,4405	0,3575	0,2851	0,2237	0,1730	0,1321	0,0996	0,0744	0,0550	0,0403	0,0293	0,0211	0,0151	0,0107	0,0076	4
5	0,7029	0,6160	0,5289	0,4457	0,3690	0,3007	0,2414	0,1912	0,1496	0,1157	0,0885	0,0671	0,0504	0,0375	0,0277	0,0203	5
6	0,8311	0,7622	0,6860	0,6063	0,5265	0,4497	0,3782	0,3134	0,2562	0,2068	0,1649	0,1301	0,1016	0,0786	0,0603	0,0458	6
7	0,9134	0,8666	0,8095	0,7440	0,6728	0,5987	0,5246	0,4530	0,3856	0,3239	0,2687	0,2202	0,1785	0,1432	0,1137	0,0895	7
8	0,9597	0,9319	0,8944	0,8472	0,7916	0,7291	0,6620	0,5925	0,5231	0,4557	0,3918	0,3328	0,2794	0,2320	0,1906	0,1550	8
9	0,9829	0,9682	0,9462	0,9161	0,8774	0,8305	0,7764	0,7166	0,6530	0,5874	0,5218	0,4579	0,3971	0,3405	0,2888	0,2424	9
10	0,9933	0,9863	0,9747	0,9574	0,9332	0,9015	0,8622	0,8159	0,7634	0,7060	0,6453	0,5830	0,5207	0,4599	0,4017	0,3472	10
11	0,9976	0,9945	0,9890	0,9799	0,9661	0,9467	0,9208	0,8881	0,8487	0,8030	0,7520	0,6968	0,6387	0,5793	0,5198	0,4616	11
12	0,9992	0,9980	0,9955	0,9912	0,9840	0,9730	0,9573	0,9362	0,9091	0,8758	0,8364	0,7916	0,7420	0,6887	0,6329	0,5760	12
13	0,9997	0,9993	0,9983	0,9964	0,9929	0,9872	0,9784	0,9658	0,9486	0,9261	0,8981	0,8645	0,8253	0,7813	0,7330	0,6815	13
14	0,9999	0,9998	0,9994	0,9986	0,9970	0,9943	0,9897	0,9827	0,9726	0,9585	0,9400	0,9165	0,8879	0,8540	0,8153	0,7720	14
15	1,0000	0,9999	0,9998	0,9995	0,9988	0,9976	0,9954	0,9918	0,9862	0,9780	0,9665	0,9513	0,9317	0,9074	0,8783	0,8444	15
16		1,0000	0,9999	0,9998	0,9996	0,9990	0,9980	0,9963	0,9934	0,9889	0,9823	0,9730	0,9604	0,9441	0,9236	0,8987	16
17			1,0000	0,9999	0,9998	0,9996	0,9992	0,9984	0,9970	0,9947	0,9911	0,9857	0,9781	0,9678	0,9542	0,9370	17
18				1,0000	0,9999	0,9999	0,9997	0,9993	0,9987	0,9976	0,9957	0,9928	0,9885	0,9823	0,9738	0,9626	18
19					1,0000	1,0000	0,9999	0,9997	0,9995	0,9989	0,9980	0,9965	0,9942	0,9907	0,9857	0,9787	19
20							1,0000	0,9999	0,9998	0,9996	0,9991	0,9984	0,9972	0,9953	0,9925	0,9884	20
21								1,0000	0,9999	0,9998	0,9996	0,9993	0,9987	0,9977	0,9962	0,9939	21
22									1,0000	0,9999	0,9999	0,9997	0,9994	0,9990	0,9982	0,9970	22
23										1,0000	0,9999	0,9999	0,9998	0,9995	0,9992	0,9985	23
24											1,0000	1,0000	0,9999	0,9998	0,9996	0,9993	24
25													1,0000	0,9999	0,9998	0,9997	25
26														1,0000	0,9999	0,9999	26
27															1,0000	0,9999	27
28																1,0000	28

Ejemplo: Sea X una P (λ = 7); entonces P(X ≤ 4) = 0,1730

Proceso de Poisson

Una aplicación importante de la distribución de Poisson se presenta en relación con el acontecimiento de eventos de un tipo particular en el tiempo. Por ejemplo, un evento podría ser un individuo entrando en un establecimiento en particular, o pulsos radiactivos registrados por un contador Geiger, o automóviles pasando por un cruce determinado. Supongamos que tenemos eventos que ocurren en ciertos puntos aleatorios de tiempo, y asumimos que para alguna constante positiva λ las siguientes suposiciones se sostienen:

1- La probabilidad que exactamente 1 evento ocurra en un intervalo de longitud t es la misma para todos los intervalos de longitud t y es igual a $\lambda t + o(t)$ (donde $o(t)$ simboliza una función $f(t)$ tal que $\lim_{t \to 0} \frac{f(t)}{t} = 0$, por ejemplo $f(t) = t^2$ es $o(t)$, pero $f(t) = t$ no lo es)

2- La probabilidad que 2 o mas eventos ocurran en un intervalo de longitud t es la misma para todos los intervalos de longitud t y es igual a $o(t)$.

3- Para cualesquiera enteros $n, k_1, k_2, ..., k_n$ y cualquier conjunto de n intervalos $I_1, I_2, ..., I_n$ que no se superpongan, si definimos los eventos E_i :"en el intervalo I_i ocurren exactamente k_i eventos" $i = 1,2,...,n$, entonces los eventos $E_1, E_2, ..., E_n$ son independientes.

3.5 Variables aleatorias continuas

Las variables aleatorias continuas son aquellas cuyo rango son todos los números reales de un intervalo dado, (es decir es un conjunto infinito no numerable). Ejemplos de variables continuas podrían ser:

X: "tiempo que tarda en llegar un colectivo a una parada"

Y: "tiempo de vida de un fusible"

Como ahora los valores de una v.a. continua no son contables no se puede hablar del i-ésimo valor de la v.a. X y por lo tanto p (x_i) = P (X=x_i) pierde su significado. Lo que se hace es sustituir la función p (x) definida sólo para x_1, x_2,.... , por una función f (x) definida para todos los valores x del rango de X. Por lo tanto se da la siguiente definición de v.a. continua

Sea X una v.a.. Decimos que es continua si existe una función no negativa f , definida sobre todos los reales x ∈ (− ∞ , ∞) , tal que para cualquier conjunto B de números reales.

$$P(X \in B) = \int_B f(x)dx$$

O sea que la probabilidad de que X tome valores en B se obtiene al integrar la función f sobre el conjunto B.

A la función f la llamamos función densidad de probabilidad (f.d.p.).

Función de distribución acumulada

Sea X una v.a. continua. Se define la función de distribución acumulada de X (abreviamos F.d.a de X) como

$$F(x) = P(X \leq x) \qquad -\infty < x < \infty$$

Si X tiene f.d.p. f(x) entonces

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \qquad -\infty < x < \infty$$

Además

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = F(b) - F(a)$$

Observaciones:

1- Si X es una v.a. con f.d.p. f(x) y función de distribución acumulada F (x) entonces

$$\frac{dF(x)}{dx} = \frac{d}{dx}\left(\int_{-\infty}^{x} f(t)dt\right) = f(x) \quad \text{donde} \quad F(x) \text{ sea derivable}$$

Es decir, se puede obtener la función de densidad de X a partir de su F.d.a.

3.6Esperanza de una variable aleatoria continua

Para una v.a. discreta la E (X) se definió como la suma de los x_i p (x_i) . Si X es una v.a. continua con f.d.p. f(x), se define E (X) sustituyendo la sumatoria por integración y p(x_i) por f(x).

La esperanza de una v.a. continua X con f.d.p. f(x) se define como

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

A menudo se desea calcular la esperanza de una función de X, Y = h(X), esto se puede hacer hallando previamente la densidad de Y y luego calcular E(Y) aplicando la definición anterior.

Otra forma de calcular E(Y) sin hallar la densidad de Y está dada por el siguiente

Teorema: Si X es una v.a. continua con f.d.p. f(x) y h(X) es cualquier función de X, entonces

$$E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$$

Varianza de una variable aleatoria continua

Sea X una v.a. continua con *f.d.p.* $f(x)$ y sea $E(X) = \mu$, entonces la varianza de X es

$$V(X) = \sigma_X^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

La interpretación de la varianza de una v.a. continua es la misma que para el caso discreto.Además sigue valiendo la igualdad

$$V(X) = E\left(X^2\right) - \mu^2$$

3.7 - Variables aleatorias continuas importantes

Distribución uniforme

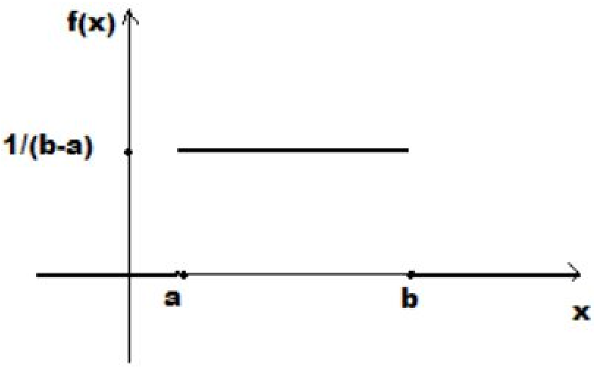
La distribución o modelo uniforme puede considerarse como proveniente de un proceso de extracción aleatoria .El planteamiento radica en el hecho de que la probabilidad se distribuye uniformemente a lo largo de un intervalo . La distribución continua más sencilla es análoga a su contraparte discreta. Una v.a. continua X se dice que tiene distribución uniforme en el intervalo a, b , con a < b, si tiene función de densidad de probabilidad dada por

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si} \quad a \le x \le b \\ 0 & \text{caso contrario} \end{cases}$$

La figura muestra la gráfica de la *f.d.p.*

<u>Notación</u>: $X \sim U[a,b]$

Es fácil verificar que $f(x)$ es una *f.d.p.* pues $f(x) \ge 0$ para todo x, y además



La F.d.a. para $a \le x \le b$ sería

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt = \int_{a}^{x} \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

Esperanza y varianza

Sea una X variable aleatoria continua distribuida uniformemente en el intervalo a , b , es decir

$$X \sim U[a,b]. \text{ Entonces, } E(X) = \frac{a+b}{2} \text{ y } \quad V(X) = \frac{(b-a)^2}{12}.$$

Distribución normal o gaussiana

La **distribución normal o gaussiana** es acampanada, simétrica, unimodal, con media y mediana similares y las colas cercanas a 0. La ventaja de la distribución gaussiana es que permite inferir la probabilidad de que se presente determinado valor y los mayores o menores que él. Ello se debe a que tiene **parámetros**, valores que nos permiten ubicarnos dentro de la distribución. Dichos parámetros son la media y el desvío estándar.

Sea X una v.a. Decimos que tiene distribución normal con parámetros y si su f.d.p. es de la forma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

Donde $\mu \in R$ y $\sigma > 0$

Notación: $X \sim N(\mu, \sigma^2)$

Observación: Cuando μ varía la gráfica de la función se traslada, μ es un parámetro de posición. Cuando σ aumenta, la gráfica se “achata”, cuando σ disminuye la gráfica se hace más “puntiaguda”,se dice que es un parámetro de escala.

Si μ =0 y σ=1 entonces se dice que X tiene distribución normal estándar. Se anota X ~ N(0,1)

En este caso la *f.d.p.* se simboliza con $\varphi(x)$, es decir

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad -\infty < x < \infty$$

En este caso la gráfica de la densidad es simétrica con respecto al origen.
La *F.d.a.* de una v.a. normal estándar se anota $\Phi(x)$

$$\Phi(x) = P(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \, dt$$

Esta integral no puede expresarse en términos de funciones elementales, por lo tanto se calcula $\Phi(x)$

Esperanza y varianza de una variable aleatoria con distribución normal

$$\boxed{\text{Sea } X \sim N\left(\mu, \sigma^2\right) \text{ entonces } E(X) = \mu \text{ y } V(X) = \sigma^2}$$

Distribución exponencial

Sea X una v.a. continua. Se dice que tiene distribución exponencial con parámetro λ si su f.d.p. es de la forma:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \ge 0 \\ 0 & \text{caso contrario} \end{cases} \qquad \text{donde } \lambda > 0$$

La distribución exponencial se utiliza algunas veces para modelar el tiempo que transcurre antes de que ocurra un evento. A menudo se lo llama tiempo de espera.

Esperanza y varianza de la distribución exponencial

$$\boxed{\text{Sea } X \sim Exp(\lambda) \text{ entonces } E(X) = \frac{1}{\lambda} \text{ y } V(X) = \frac{1}{\lambda^2}}$$

Propiedades de la distribución exponencial

1- Relación entre la distribución exponencial y el proceso de Poisson.

Sea T el tiempo de espera hasta el siguiente evento en un proceso de Poisson con parámetro λ.
Veamos cuál es la $F.d.a.$ de T.
Si $t < 0$ entonces claramente $F(t) = P(T \le t) = 0$
Si $t \ge 0$ entonces para hallar $F(t) = P(T \le t)$ consideramos el evento complementario de $\{T \le t\}$.
Notar que $\{T > t\}$ si y solo si no ocurre ningún evento durante las siguientes t unidades de tiempo.
Si X: "número de eventos que ocurren en las siguientes t unidades de tiempo", entonces $\{T > t\}$ ocurre si y solo si $\{X = 0\}$ ocurre, por lo tanto $P(T > t) = P(X = 0)$
Como $X \sim P(\lambda t)$ entonces

$$P(T > t) = P(X = 0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

Por lo tanto

$$F(t) = P(T \le t) = 1 - P(T > t) = 1 - e^{-\lambda t}$$

Como $F(t)$ es la F.d.a. de una v.a. exponencial, entonces $T \sim Exp(\lambda)$
Por lo tanto

> Si los eventos siguen un proceso de Poisson con parámetro λ, y si T representa el tiempo de espera desde cualquier punto inicial hasta el próximo evento, entonces $T \sim Exp(\lambda)$

2- Propiedad falta de memoria

La distribución exponencial tiene una propiedad conocida como falta de memoria, que se muestra en el siguiente ejemplo:
El tiempo de vida, en años, de un circuito integrado particular tiene una distribución exponencial con parámetro 0.5. Encuentre la probabilidad de que el circuito dure más de tres años. Sea la v.a. X : "tiempo de vida, en años, de un circuito integrado particular", entonces X~Exp(0.5)

$$\text{Y } P(X > 3) = 1 - F(3) = e^{-1.5} = 0.223$$

Supongamos ahora que actualmente un circuito tiene cuatro años y aún funciona. Se quiere hallar la probabilidad de que funcione tres años más. Por lo tanto planteamos una probabilidad condicional

$$P/X > 7/X > 4) = \frac{P(X > 7 \text{ y } X > 4)}{P(X > 4)} = \frac{P(X > 7)}{P(X > 4)} = \frac{e^{-0.5\times 7}}{e^{-0.5\times 4}} = e^{-0.5\times(7-4)} = e^{-1.5} = 0.223$$

Observamos que $P/X > 7/X > 4) = P(X > 7-4) = P(X > 3)$

En general, la probabilidad que se tenga que esperar t unidades adicionales, dado que ya se han esperado s unidades, es la misma que la probabilidad de que se tenga que esperar t unidades desde el inicio. **La distribución exponencial no "recuerda" cuánto tiempo se**

ha esperado. En particular, si el tiempo de vida de un componente sigue una distribución exponencial, entonces la probabilidad de que un componente que tiene s unidades de tiempo dure t unidades de tiempo adicionales es la misma que la probabilidad de que un componente nuevo dure t unidades de tiempo. En otras palabras, un componente cuyo tiempo de vida siga una distribución exponencial no muestra ningún síntoma de los años o del uso. Los cálculos hechos en el ejemplo anterior se pueden repetir para valores cualesquiera s y t y entonces
se pudo probar que

si $X \sim Exp(\lambda)$ y t y s son números positivos, entonces $P(X > t + s / X > s) = P(X > t)$

4 - DESIGUALDAD DE CHEBYSHEV- LEY DE LOS GRANDES NÚMEROS

La desigualdad de Chebyshev es una importante herramienta teórica. Entre otras aplicaciones constituirá un medio para comprender cómo la varianza mide la variabilidad de una dada variable aleatoria, con respecto a su esperanza matemática. También nos permitirá establecer con más precisión el hecho, reiteradamente señalando, de que la frecuencia relativa f_A de un suceso A asociado a un experimento aleatorio ε tiende, cuando el número de repeticiones de ε se hace infinitamente grande, a la probabilidad P(A) (resultado conocido como la Ley de los grandes números). Pero además es de utilidad práctica pues, al constituir una cota de ciertas probabilidades, nos podrá servir como una estimación de esas mismas probabilidades.

4.1-Desigualdad de Chebyshe

Sea X una variable aleatoria cuya esperanza es E(X) , sea c un número real cualquiera y supongamos que

$E\left[(X-c)^2\right]$ existe y es finito. Entonces

$$\boxed{\forall \varepsilon > 0 \qquad P\left(|X-c| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} E\left[(X-c)^2\right]}$$

Formas alternativas de la desigualdad de Chebyshev.

Podemos escribir la desigualdad de Chebyshev en una serie de formas alternativas:

a_1) Tenemos en primer lugar la forma que acabamos de demostrar:

$$P(|X-c| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E\left[(X-c)^2\right]$$

a_2) Si consideramos el suceso complementario a $|X-c| \geq \varepsilon$, es decir $|X-c| < \varepsilon$, podemos escribir, recordando que $P(\overline{A}) = 1 - P(A)$:

$P(|X-c| < \varepsilon) = 1 - P(|X-c| \geq \varepsilon) \geq 1 - \frac{1}{\varepsilon^2} E\left[(X-c)^2\right]$, esto es:

$$P(|X-c| < \varepsilon) \geq 1 - \frac{1}{\varepsilon^2} E\left[(X-c)^2\right]$$

b_1) Si en a_1) elegimos $c = E(X)$ tenemos $P(|X - E(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E\left([X - E(X)]^2\right)$ y recordando la definición de la varianza: $V(X) = E\left([X - E(X)]^2\right)$ tenemos:

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

b_2) La correspondiente expresión para el suceso complementario es:

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}$$

c_1) Si en a_1) elegimos $c = E(X)$ y además elegimos $\varepsilon = k\sigma_X = k\sqrt{V(X)}$, tenemos $P(|X - E(X)| \geq k\sigma_X) \leq \frac{V(X)}{(k\sigma_X)^2} = \frac{\sigma_X^2}{(k\sigma_X)^2}$, es decir:

La varianza como una medida de la concentración de la fdp de una v.a. alrededor de la esperanza.

Podemos usar las formas b) de la desigualdad de Chebyshev para interpretar a la varianza V(X) como una medida de la variabilidad de la variable aleatoria X con respecto a su esperanza o en otras palabras de cómo la distribución de la v.a. X se concentra o dispersa con respecto a la esperanza E(X).

De la expresión $P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$ vemos que, para un ε dado, si $V(X)$ es muy pequeño entonces la probabilidad de que X tome valores lejos de E(X) es muy chica, es decir hay una gran probabilidad de que X tome valores próximos a E(X) . Inversamente si V(X) es grande, la probabilidad de que X tome valores alejados de E(X) puede ser también grande. Podemos precisar un poco más esto considerando el siguiente

Teorema. Si $V(X) = 0$ entonces $P[X = E(X)] = 1$. Decimos que $X = E(X)$ con probabilidad 1 (X es igual a su esperanza con probabilidad 1).

4.2 - La ley de los grandes números.

La ley de los grandes números establece en forma precisa el hecho que cuando el número de repeticiones de un experimento se hace muy grande, la frecuencia relativa f_A de un suceso A relacionado con el experimento converge en sentido probabilístico a la probabilidad P(A). Daremos una versión de la Ley de los grandes números conocida como la forma de Bernoulli .

Teorema (Forma de Bernoulli de la ley de los grandes números)
Sea un experimento probabilístico y sea A un suceso asociado con él. Consideremos n repeticiones independientes del experimento. Sea n_A el número de veces que ocurre A en las n repeticiones de forma tal que $f_A = \frac{n_A}{n}$ es la frecuencia relativa. Sea $P(A) = p$ (que se supone igual para todas las repeticiones). Entonces, para cualquier número $\varepsilon > 0$ se cumple

$$P(|f_A - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

5- VARIABLES ALEATORIAS BIDIMENSIONALES

5.1 – Generalidades

Hasta ahora hemos considerado el caso de variables aleatorias unidimensionales. Esto es, el resultado del experimento de interés se registra como un único número real. En muchos casos, sin embargo, nos puede interesar asociar a cada resultado de un experimento aleatorio, dos o más características numéricas. Por ejemplo, de los remaches que salen de una línea de producción nos puede interesar el diametro X y la longitud Y. Teniendo en cuenta la inevitable variabilidad en las dimensiones de los remaches debido a las numerosas causas presentes en el proceso de fabricación, los podemos representar asociándose dos variables aleatorias X e Y que pueden pensarse como una variable aleatoria bidimensional: (X , Y) .

Al conjunto de valores que toma la variable aleatoria bidimensional (X,Y) lo llamaremos ***recorrido*** de la v.a. (X,Y) y lo indicaremos R_{XY}. En otras palabras $R_{XY} = \left\{(x,y) : x = X(s) \ \ e \ \ y = Y(s) \ \ con \ \ s \in S\right\}$, es decir, es la imagen por (X,Y) del espacio muestral S.

Clasificaremos a las variables aleatorias bidimensionales de la siguiente manera:
(X,Y) es **v.a. bidimensional discreta** si X e Y son discretas

(X,Y) es **v.a. bidimensional continua** si X e Y son continuas
El caso X continua, Y discreta (o viceversa) no lo consideramos.
Sea (X,Y) una variable aleatoria bidimensional discreta y sea R_{XY} su recorrido (numerable). Sea $p : R_{XY} \to R$ una función que a cada elemento (x_i, y_j) le asigna un número real $p(x_i, y_j)$ tal que

$$P\left(X = x_i,\ Y = y_j \right) = p(x_i, y_j) \quad \forall (x_i, y_j) \in R_{XY}$$ y que verifica.

a) $p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$

b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = \sum_i \sum_j p(x_i, y_j) = 1$

A esta función la llamaremos ***función de probabilidad puntual conjunta*** de la variable aleatoria bidi mensional (X,Y). En forma abreviada la designaremos *fdp conjunta*.

Ejemplo 1: Dos líneas de producción, señaladas I y II, manufacturan cierto tipo de artículo a pequeña escala. Supóngase que la capacidad máxima de producción de la línea I es cinco artículos por día, mientras que para la línea II es 3 artículos/día. Debido a los innumerables factores presentes en todo proceso de producción, el número de artículos realmente producido por cada línea puede pensarse como una variable aleatoria. En conjunto podemos pensar en una variable aleatoria bidimensional (X , Y) discreta, donde la primera componente X corresponde a la producción de la línea I y la segunda componente Y a los artí-
culos que salen de la línea II. La fdp conjunta correspondiente a variables aleatorias bidimensionales suele presentarse, por comodidad, como una tabla. Supongamos que la para la v.a. (X , Y) que nos interesa aquí la tabla correspondiente a p (x_i , y_j) es

Y/X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

¿Cuál es la probabilidad de que salgan más artículos de la línea I que de la línea II?
Antes de calcular la probabilidad que nos pide el problema, hagamos algunas consideraciones sobre la tabla que representa a p (x_i , y_j).

5.2 - Funciones de distribución marginales de una v.a. (X,Y) discreta