

Modelos

- Probabilidad y variables aleatorias
 - Probabilidad y sus propiedades
 - Probabilidad condicionada
 - Variables aleatorias
- Modelos de distribución de probabilidad
 - El proceso de Bernoulli
 - El proceso de Poisson
 - Las distribuciones de duraciones de vida
 - La distribución Normal
 - La distribución Log-Normal
- Modelos Multivariantes

Modelos: Probabilidad y variables aleatorias

Probabilidad y variables aleatorias

- Probabilidad y sus propiedades
- Probabilidad condicionada
- Variables aleatorias

Probabilidad y sus propiedades: Concepto

- Cuando tenemos una muestra de una población, después de describirla, nuestro objetivo es inferir las propiedades de la población a partir de la muestra.
- El instrumento conceptual que permite la generalización es un **modelo de la población**.
- El calculo de probabilidades me permite calcular este modelo que actúa de puente entre lo observado, i.e. la muestra, y lo desconocido, i.e. la población.
- La probabilidad de una población finita homogénea de N elementos, k de los cuales tienen la característica A en una primera aproximación se puede poner como:
 - $P(A)=k/N$

Probabilidad y sus propiedades: Concepto



Probabilidad y sus propiedades: Concepto

- Hay que ser consciente que muchas veces no se puede obtener un conocimiento exacto de la probabilidad ya que:
 - Al no ser posible una experimentación definida, siempre tenemos una información limitada sobre la frecuencia relativa.
 - El sistema observado puede variar a lo largo del tiempo, y por tanto también las frecuencias relativas.
- Por tanto para poblaciones finitas la identificación de la probabilidad con la frecuencia relativa es casi inmediata, pero para poblaciones infinitas puede presentar serios problemas.
- Esto se complica mucho más para sucesos inciertos, que solamente ocurren un número de veces muy reducido.

Probabilidad y sus propiedades: Definición

- Población: conjunto de elementos homogéneos en los que se desea investigar la ocurrencia de una característica o propiedad:
 - Numero de elementos finito o infinito.
 - Debe ser posible observar sus elementos.
 - Debe ser posible saber si un elemento pertenece a ella o no.
- Sucesos elementales: es el conjunto de resultados posibles de una característica o propiedad que se investiga, que verifican:
 - Siempre ocurre alguno de ellos.
 - Son mutuamente excluyentes.
- Sucesos compuestos: los contruidos a partir de uniones de resultados elementales.
- P. ej. Tirar un dado:
 - Sucesos elementales: 1, 2, 3, 4, 5, 6
 - Sucesos compuestos: número par, número impar, menor 5, múltiplo de 2, etc.

Probabilidad y sus propiedades: Definición

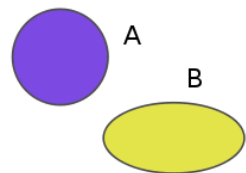
- Espacio muestral: conjunto de todos los resultados posibles del experimento. Así los elementos elementales y compuestos son los subconjuntos del espacio muestral.
 - Definimos el suceso seguro, E , como todo el espacio muestral: es seguro porque siempre ocurre.
 - Definimos el suceso imposible, \emptyset , como aquel que no ocurre nunca.

Probabilidad y sus propiedades: Definición

➤ Se desea asociar a cada suceso una medida de incertidumbre que llamaremos probabilidad, con las siguientes propiedades:

1. Debido a que la $f_r(A)$ es un numero comprendido entre 0 y 1, la probabilidad: $0 \leq P(A) \leq 1$.
2. La frecuencia del suceso seguro ocurre siempre: $P(E)=1$.
3. Si A y B son características mutuamente excluyentes y las unimos en una nueva $C=A+B$. C ocurre cuando ocurre A o ocurre B. La frecuencia relativa de C es la suma de las frecuencias relativas de A y B, y por tanto la probabilidad de sucesos mutuamente excluyentes:
 $P(A+B)=P(A \cup B)=P(A \text{ or } B)=P(A)+P(B)$. P. ej probabilidad de sacar 1 o 2 en una tirada de dado: $P(1 + 2)=P(1 \cup 2)=P(1 \text{ or } 2)=P(1)+P(2)=1/6+1/6=1/3$.

Imagen extraída de
https://es.wikipedia.org/wiki/Probabilidad_condicionada



mutuamente
excluyentes

Probabilidad y sus propiedades: Definición

- Se desea asociar a cada suceso una medida de incertidumbre que llamaremos probabilidad, con las siguientes propiedades:
- 4. Suponiendo **A y B NO son mutuamente excluyentes en general**, podemos definir n_{AB} , $n_{A\bar{B}}$, $n_{\bar{A}B}$ como el numero de veces que aparecen los 3 sucesos compuestos mutuamente excluyentes (A y B), (A y no B), (no A y B), tendremos:
 - $n_A = n_{AB} + n_{A\bar{B}}$, $n_B = n_{AB} + n_{\bar{A}B}$, $n_{A+B} = n_{AB} + n_{A\bar{B}} + n_{\bar{A}B}$, de estas 3 ecuaciones se puede obtener $n_{A+B} = n_A + n_B - n_{AB}$ (sustituyendo $n_{\bar{A}B}$ y $n_{A\bar{B}}$)
 - Si dividimos por el número total de observaciones obtenemos las frecuencias relativas que pasadas a probabilidades:
 - $P(A+B) = P(A) + P(B) - P(AB)$
 - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 - $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
 - **P. ej.** Probabilidad de elegir una carta al azar de una baraja sea un **corazón** o una **cara** (J, Q, K) es $13/52 + 12/52 - 3/52 = 11/26$ (de 52 cartas 13 son corazones, 12 caras , y 3 son corazones y caras).

Probabilidad y sus propiedades: Definición

- Se desea asociar a cada suceso una medida de incertidumbre que llamaremos probabilidad, con las siguientes propiedades:
 - 5. Suponiendo que \bar{A} es el suceso complementario de A, que ocurre siempre que no ocurre A, de las propiedades anteriores se puede deducir:
 - $P(\bar{A}) = 1 - P(A)$.

Probabilidad y sus propiedades: Definición

- **Detalle de A y B NO mutuamente excluyentes:**
- En particular con dos sucesos no mutuamente excluyentes, podemos definir 4 sucesos siguientes:
 - El suceso (A y B) que aparecerá n_{AB} veces.
 - El suceso (A y no B) que aparecerá $n_{A\bar{B}}$ veces.
 - Y el suceso (no A y B) que aparecerá $n_{\bar{A}B}$ veces.
 - El suceso (no A y no B) que aparecerá 0 veces.
 - En cuanto al número de sucesos podemos poner:
 - $n_A = n_{AB} + n_{A\bar{B}}$: el número de veces que ocurre A es igual al número de veces que ocurre A y B a la vez más el número de veces que ocurre A y no ocurre B.
 - $n_B = n_{AB} + n_{\bar{A}B}$: el número de veces que ocurre B es igual al número de veces que ocurre B y A a la vez más el número de veces que ocurre B y no ocurre A.
 - $n_{A+B} = n_{AB} + n_{A\bar{B}} + n_{\bar{A}B}$: el número de veces que ocurre A o ocurre B es igual al número de veces que ocurre A y B a la vez, más el número de veces que ocurre A y no ocurre B, más el número de veces que no ocurre A y ocurre B. Estas son todas la posibilidades.
 - De estas 3 ecuaciones se puede obtener $n_{A+B} = n_A + n_B - n_{AB}$ (sustituyendo $n_{\bar{A}B}$ y $n_{A\bar{B}}$)

Probabilidad Condicionada

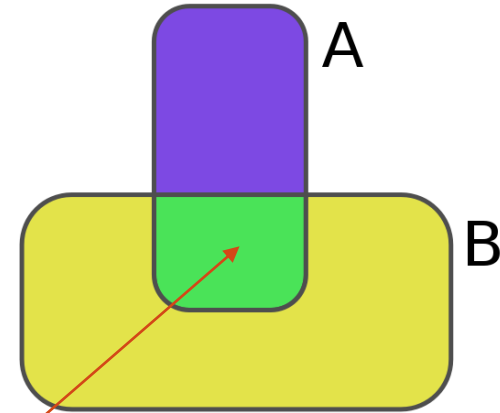
- La frecuencia relativa de A condicionada a la ocurrencia de B se define considerando los casos en los que aparece B, y viendo en cuantos de estos casos ocurre A (casos que aparecen A y B, dividido por el numero de casos que aparece B):

- $f_r(A | B) = n_{AB} / n_B$, como $f_r(A) = n_A / n$, $f_r(B) = n_B / n$ y $f_r(AB) = n_{AB} / n$, se tiene:
- $f_r(A | B) = f_r(AB) / f_r(B)$ o lo que es lo mismo:
- $f_r(AB) = f_r(A | B) f_r(B) = f_r(B | A) f_r(A)$

- Así exigiremos la misma propiedad a la probabilidad y definiremos probabilidad de un suceso A condicionada a otro B como:

- $P(A | B) = P(AB) / P(B)$, ojo es una intersección en $P(B)$

Imagen extraída de https://es.wikipedia.org/wiki/Probabilidad_condicionada



Tomando los casos en los que B se cumple, la fracción en los que también se cumple A.

Probabilidad Condicionada: Independencia de sucesos

- Diremos que A y B son sucesos independientes, si el conocimiento de la ocurrencia de uno no modifica la probabilidad de aparición del otro:
 - $P(A | B) = P(A)$ o $P(B | A) = P(B)$, o lo que es lo mismo:
 - $P(AB) = P(A) P(B)$
- Esto se puede generalizar a cualquier numero de sucesos:
 - Diremos que los sucesos A_1, \dots, A_n , son independientes si la probabilidad conjunta de cualquier subconjunto que pueda formarse con ellos es el producto de las probabilidades individuales.

Probabilidad Condicionada: Teorema de Bayes

- Consideremos un experimento que se realiza en dos etapas.
- **En la primera** los sucesos posibles, A_1, \dots, A_n , son mutuamente excluyentes con probabilidades conocidas, $P(A_i)$ y con $\sum P(A_i)=1$.
- La $P(A_i)$ son las **probabilidades a priori**.
- **En la segunda** los resultados posibles, B_i , dependen de los de la primera y se conocen la probabilidades condicionadas $P(B_i | A_i)$, de obtener cada posible resultado B_i , cuando aparece en la primera etapa el suceso A_i .
- Se efectúa ahora el experimento, pero el resultado de la primera fase A_i , no se conoce, pero si el de la segunda fase que es B_i .
- Las $P(A_i | B_i)$ son las **probabilidades a posteriori**, y es lo que calcula el Teorema de Bayes.

Probabilidad Condicionada: Teorema de Bayes

- El teorema de Bayes permite calcular la probabilidades $P(A_i | B_i)$, de los sucesos no observados en la primera etapa, dado el resultado observado en la segunda (**probabilidades a posteriori**).
- Se calcula partiendo de la definición de probabilidad condicionada:
 - $P(A_i | B_i) = P(A_i B_i) / P(B_i) = P(B_i | A_i) P(A_i) / P(B_i)$
- Por otro lado $P(B_i) = P(B_i A_1 + B_i A_2 + \dots + B_i A_n)$, ya que B_i debe ocurrir con alguno de los sucesos de A_i .
- Como los sucesos $B_i A_1, B_i A_2, \dots$ son mutuamente excluyentes ya que por la definición del problema los A_i lo son también, entonces:
 - $P(B_i) = \sum_i P(B_i A_i) = \sum_i P(B_i | A_i) P(A_i)$
- Si sustituimos arriba obtenemos el Teorema de Bayes:
 - $P(A_i | B_i) = P(B_i | A_i) P(A_i) / P(B_i) = P(B_i | A_i) P(A_i) / \sum_i P(B_i | A_i) P(A_i)$

[Volver](#)
[Estimación](#)
[Bayesiana](#)

Probabilidad Condicionada: Teorema de Bayes

- Dos urnas: U_1 (70% b, 30% n) y U_2 (30% b, 70% n): Sucesos A_1 y A_2
- Se selecciona una urna al azar y se saca el resultado de bolas con remplazamiento: bnbnnnnnnnn. Es el suceso B,
- ¿Cuál es la probabilidad de que la muestra venga de U_1 ? Es decir $P(U_1 | B)$???
- Se puede considerar el experimento en dos etapas:
 - Selección de urna.
 - Extracción de la muestra de la urna seleccionada.
- Selección al azar de las urnas $P(U_1)=P(U_2)=0.5$
- Los 10 sucesos de sacar bolas son independientes (extracción con remplazamiento): $P(b | U_1)=0.7$, $P(n | U_1)=0.3$ y $P(b | U_2)=0.3$, $P(n | U_2)=0.7$ (las probabilidades dependen de la urna)
- Se verifica $P(B | U_1)=P(bnbnnnnnnnn | U_1)=P(b | U_1) P(n | U_1) P(b | U_1) \dots P(b | U_1)=0.7^8 0.3^2$
- Se verifica $P(B | U_2)=P(bnbnnnnnnnn | U_2)=P(b | U_2) P(n | U_2) P(b | U_2) \dots P(b | U_2)=0.3^8 0.7^2$
- La probabilidad pedida es $P(U_1 | B)$ que la calculamos por el Teorema de Bayes:
 - $P(U_1 | B)=P(B | U_1)P(U_1)/\sum_i P(B | U_i) P(U_i)=P(B | U_1)P(U_1)/(P(B | U_1) P(U_1)+P(B | U_2) P(U_2))$
 - $P(U_1 | B)=(0.7^8 0.3^2 0.5)/(0.7^8 0.3^2 0.5+ 0.3^8 0.7^2 0.5)=0.994$ (para casa calcular $P(U_2 | B)$)

Variables aleatorias: Discretas

- El cálculo de probabilidades utiliza variables numéricas que se denominan aleatorias, ya que sus valores se determinan al azar.
- Variables aleatorias discretas: cuando toma un número de valores finitos o infinito numerable.
- Función de probabilidad: una variable discreta aleatoria se define por sus posibles valores discretos (su espacio muestral), junto con sus probabilidades respectivas. Así la función de probabilidad $p(x)$ es la función que indica las probabilidades de cada posible valor:
 - $p(x_i)=P(x=x_i)$, con $\sum_i p(x_i)=1$.

Variables aleatorias: Discretas

- Función de distribución: $F(x_0)=P(x \leq x_0)$ (o Función de Distribución Acumulada)
- Si suponemos que la variable x toma los posibles valores:
 - $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$
- La función distribución viene determinada por:
 - $F(x_1)=P(x \leq x_1)=p(x_1),$
 - $F(x_2)=P(x \leq x_2)=p(x_1)+p(x_2),$
 - $F(x_3)=P(x \leq x_3)=p(x_1)+p(x_2)+p(x_3),$
 -
 - $F(x_n)=P(x \leq x_n)= \sum_i p(x_i)=1.$

Variables aleatorias: Continuas

- Una variable aleatoria es continua, cuando puede tomar cualquier valor en un intervalo.
- Función de densidad: es una función continua que verifica las condiciones:
 - $f(x) \geq 0$.
 - $\int_{-\infty}^{+\infty} f(x)dx = 1$.
- El conocimiento de la función de densidad $f(x)$ nos permite calcular las probabilidades a nuestra conveniencia:
 - $P(x \leq x_0) = \int_{-\infty}^{x_0} f(x)dx$.
 - $P(x_0 \leq x \leq x_1) = \int_{x_0}^{x_1} f(x)dx$.

Variables aleatorias: Continuas

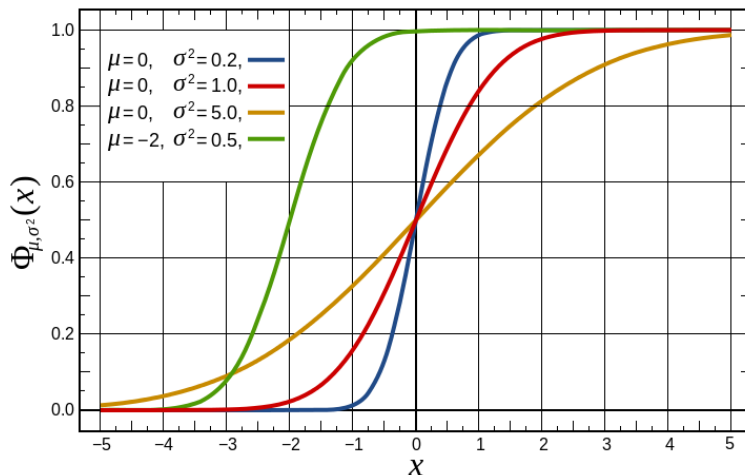
➤ La función de distribución

➤ $F(x_0) = P(x \leq x_0) = \int_{-\infty}^{x_0} f(x)dx.$

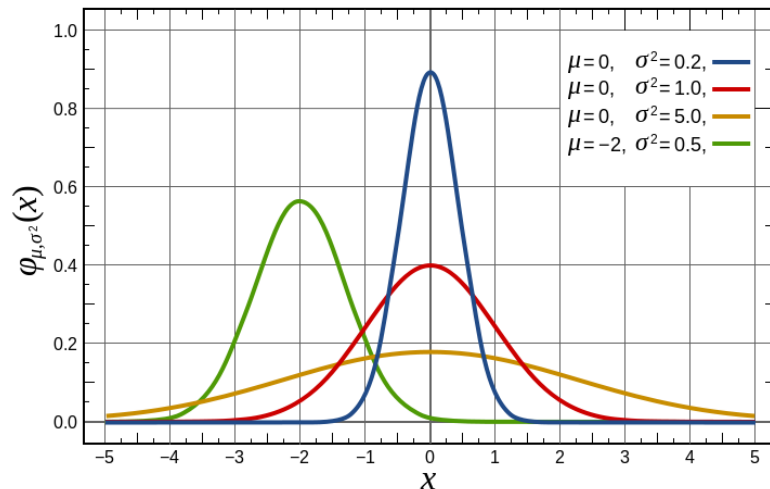
➤ $f(x) = \frac{dF(x)}{dx}$

➤ Ejemplos para la distribución normal

Función de Distribución



Función de Densidad



Variables aleatorias: Medidas de Centralización

- La mas utilizada es la media (μ) o esperanza matemática $E(x)$:
 - $\mu = E(x) = \sum_i x_i p(x_i)$ (caso discreto).
 - $\mu = E(x) = \int_{-\infty}^{+\infty} x f(x) dx$ (caso continuo).
- La mediana: es aquel valor de la variable aleatoria que divide la probabilidad total en dos mitades iguales:
 - Para variables discretas se define la mediana como el menor valor x_m de la variable aleatoria que satisface:
 - $F(x_m) \geq 0.5$, entonces x_m es la mediana.
 - Para variables continuas será el valor m definido por
 - $F(m) = 0.5 = P(x \leq m)$

Variables aleatorias: Medidas de Dispersión

- La medida de dispersión asociada a la media, es la desviación típica cuyo cuadrado es la varianza:
 - $Var(x) = \sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$
- Para variables discretas las integrales se convierten en sumas y las probabilidades $p(x)$ sustituyen a los elementos de probabilidad $f(x)dx$.
- El percentil p de una variable aleatoria x discreta es el valor x_p que verifica:
 - $P(x < x_p) \leq p$ y $P(x \leq x_p) \geq p$.
- Para variables continuas:
 - $F(x_p) = p$.

Variables aleatorias: Medidas de Dispersión

- Los **cuartiles** dividen la distribución en 4 partes iguales. La mediana coincide con el segundo cuartil, y con el percentil 0.5.
- La medida absoluta de dispersión más utilizada es el **rango intercuartílico**:
 - $RIC = Q_3 - Q_1$, que representa la zona central donde se encuentra el 50% de la probabilidad.
 - Para distribuciones simétricas: $Q_2 - Q_1 = Q_3 - Q_2$ y por lo tanto RIC es el doble de la distancia entre la mediana y los cuartiles.
- La medida de dispersión que se asocia a mediana es la **Meda**:
 - $Meda = \text{Mediana}(|x - \text{Med}(x)|)$

Variables aleatorias: Medidas de Dispersión

- La métrica **Meda** ($\text{Meda} = \text{Mediana}(|x - \text{Med}(x)|)$) para **distribuciones simétricas** cumple:
 - El 50% de las desviaciones son menores que $Q_3 - \text{Med} = \text{Med} - Q_1$.
 - El otro 50% de las desviaciones son mayores que $Q_3 - \text{Med} = \text{Med} - Q_1$.
 - Esto se entiende ya que todos los valores x que cumplen $Q_1 \leq x \leq Q_3$ tienen su meda menor que el rango: $Q_3 - \text{Med} = \text{Med} - Q_1$. Y todos los x que están fuera del intervalo (Q_1, Q_3) , tienen su Meda mayor que el rango: $Q_3 - \text{Med} = \text{Med} - Q_1$.
 - En consecuencia, solo para distribuciones simétricas, la Meda se puede calcular como $\text{Meda} = Q_3 - \text{Med}(x)$, i.e. la mitad de RIC.

Variables aleatorias: Otras Medidas Características

- En general definimos momento de orden k (m_k), respecto al origen de una variable continua aleatoria, x , como (funciones generatrices, Apéndice 5A):
 - $m_k = \int x^k f(x) dx.$
- Si tomamos como origen la media μ :
 - $\mu_k = \int (x - \mu)^k f(x) dx.$
- El coeficiente de asimetría, CA , se define como:
 - $CA = \frac{\mu_3}{\sigma^3}$
- El apuntamiento o curtosis como:
 - $CA_p = \frac{\mu_4}{\sigma^4}$
- Y el coeficiente de variación como:
 - $CV = \frac{\sigma}{|\mu|}$

Variables aleatorias: Acotación de Tchebychev

- Si conocemos la media y desviación típica de una variable aleatoria discreta o continua nos permite calcular la proporción de la distribución que está situada en el rango
 - $\mu \pm k\sigma$, siendo k una constante positiva.
- Al igual que hacíamos con las frecuencias relativas la acotación de Tchebychev siempre nos permite verificar que:
 - $P(\mu - k\sigma \leq x \leq \mu + k\sigma) \geq 1 - (1/k^2)$, para cualquier valor de k .
 - Para cualquier variable aleatoria el intervalo $\mu \pm 3\sigma$, contiene al menos el 89% de la distribución.
 - Para cualquier variable aleatoria el intervalo $\mu \pm 4\sigma$, contiene al menos el 94% de la distribución.

Ejemplos para realizar en casa

- Realizar para casa los ejemplos del libro:
 - 4-1, 4-2, 4-3, 4-4, 4-5, 4-6, 4-7, 4-8, 4,9 y 4-10.

Modelos: Modelos de distribución de probabilidad

Modelos de distribución de probabilidad

- El proceso de Bernoulli
- El proceso de Poisson
- Las distribuciones de duraciones de vida
- La distribución Normal
- La distribución Log-Normal

El proceso de Bernoulli

- Supongamos un experimento donde se observan elementos de una población que cumplen las siguientes características:
 - Cada observación se puede clasificar en dos posibles categorías:
 - Aceptable (A), con probabilidad de ocurrencia $q=1-p$.
 - Defectuoso (D), con probabilidad de ocurrencia p .
 - La proporción de elementos A y D es constante a lo largo de todo el proceso de observación.
 - Las observaciones son independientes:
 - Es decir la probabilidad de ocurrencia de los sucesos A y D no se modifican por el orden en el se observan los mismos.
- Este modelo aplica a poblaciones finitas de las que tomamos elemento al azar con reemplazamiento.
- También aplica a poblaciones conceptualmente infinitas en las cuales se generan los objetos con esas dos características con p y $1-p$.

El proceso de Bernoulli: Distribución de Bernoulli

- Definimos una **variable aleatoria de Bernoulli** con espacio muestral $\{0,1\}$, por
 - $x = \begin{cases} 0 & \text{si el elemento es aceptable} \\ 1 & \text{si el elemento es defectuoso} \end{cases}$
- La función de probabilidades de esta variable, i.e. la probabilidad de que la variable aleatoria tome el valor 0 o 1 es:
 - $P(x)=p^x q^{1-x}, x=0,1$
- Su media es
 - $\mu=E(x)=0*(1-p)+1*p=p$
- La desviación típica
 - $DT(x)=\sigma=[(0-p)^2(1-p)+(1-p)^2p]^{1/2} = (pq)^{1/2}$, (recordar que la varianza está dada por $Var(x) = \sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx$ y lo aplicamos en el caso discreto).
- En esta distribución la media y la varianza dependen de p , y la varianza será máxima cuando:
 - $d[p(1-p)]/dp=1-2p=0$, o lo que es lo mismo cuando $p=0.5$.

El proceso de Bernoulli: Distribución de Bernoulli

- La variable binomial en un **proceso de Bernoulli** se define como:
 - y = número de elementos defectuosos al observar n observaciones.
 - El espacio muestral de la nueva variable y es $\in(0,n)$.
- Supongamos que hemos realizado n observaciones de las cuales r son defectuosas y $n-r$ son aceptables (da igual el orden por la hipótesis de independencia).
- La probabilidad de ocurrencia de un suceso $\overbrace{DD\dots D}^r \overbrace{AA\dots A}^{n-r}$, es:
 - $\overbrace{pp\dots p}^r \overbrace{(1-p)(1-p)\dots(1-p)}^{n-r} = p^r(1-p)^{n-r}$, (**conservación de proporción de 0's + 1's**).
- Pero de cuantas formas podemos permutar las ocurrencias de cadenas de n objetos de las cuales r son defectuosas y $(n-r)$ aceptables. Son la permutaciones de n con r y $(n-r)$ repetidos:
 - $\frac{n!}{r!(n-r)!} = \binom{n}{r}$, es decir el número combinatorio.

El proceso de Bernoulli: Distribución de Bernoulli

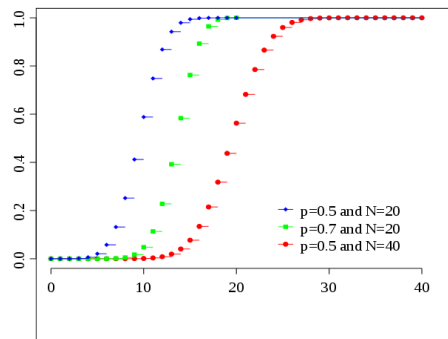
- Por tanto la probabilidad de que en una observación de n elementos r sean defectuosos viene determinado por (ver **ejemplo 5.2**):

- $$P(y = r) = \binom{n}{r} p^r (1 - p)^{n-r}, r=0, 1, \dots, n$$

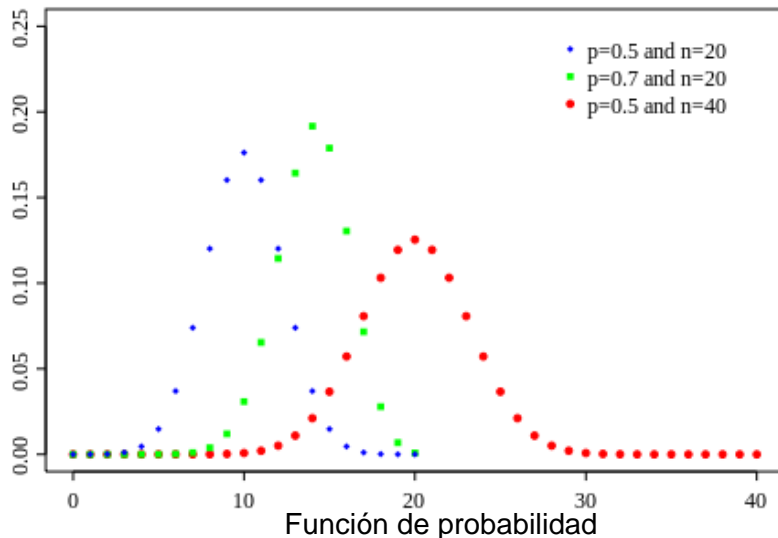
Imágenes extraídas de

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

- Es fácil comprobar que (ejercicio):
- $E[y] = \sum r P(y=r) = np$
- $DT(y) = \sigma = [\sum (r - np)^2 P(y=r)]^{1/2} = (npq)^{1/2}$.



Función de distribución de probabilidad



Función de probabilidad

El proceso de Bernoulli: Distribución geométrica

- Consideremos el mismo proceso de Bernoulli, pero en vez de contar número de defectos en n observaciones, no preguntamos por el número de elementos o numero de observaciones hasta el primer defectuoso, esto es la variable geométrica:
- x =número de elementos hasta el primer defectuoso
- Para calcular la función de probabilidad, observemos que x tomará el valor n únicamente en el suceso:

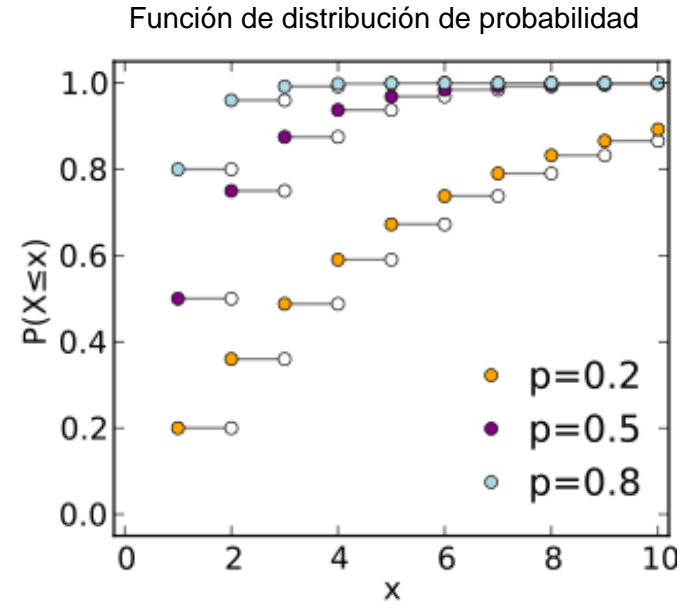
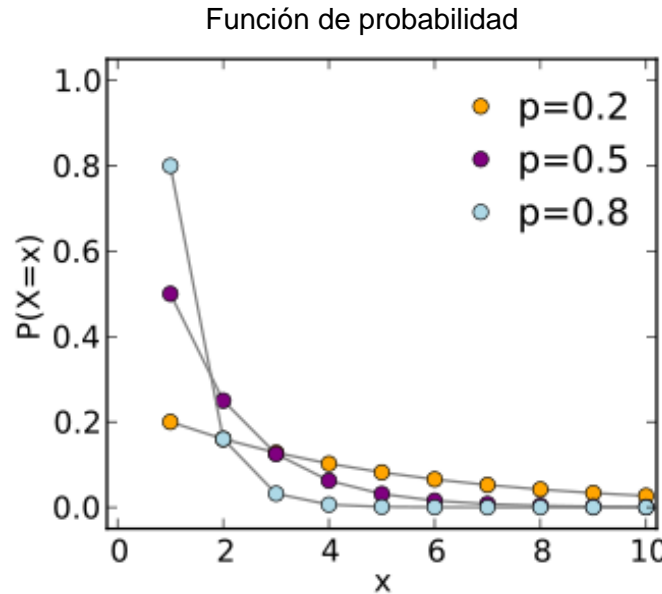
$\overbrace{AA \dots A}^{n-1} D$

- Por tanto por la independencia:
 - $P(x=n)=p(1-p)^{n-1}$, $n=1, 2, \dots$ (ver ejemplo 5.3).
 - Los parámetros de media y desviación son $E[x]=1/p$ y $Var[x]=q/p^2$
- Observar que la variable geométrica tiene un conjunto ilimitado de posibles valores, aunque su probabilidad esta normalizada:
- $\sum_1^{\infty} P(x = n) = p \sum_{n=1}^{\infty} (1-p)^{n-1} = 1$ (recordar que el termino de una progresión geométrica es $a_n = a_1 r^{n-1}$, y para una $r < 1$ $\sum_{n=1}^{\infty} a_n = \frac{a_1}{1-r}$).

El proceso de Bernoulli: Distribución geométrica

- La media y la desviación típica de la distribución geométrica se puede deducir fácilmente que es (ejercicio, apéndice 5A):
- $E[x] = 1/p$
- $Var[x] = q/p^2$

Imágenes extraídas de
https://en.wikipedia.org/wiki/Geometric_distribution



El proceso de Poisson

- Supongamos un experimento en que observamos la aparición de sucesos puntuales sobre un soporte continuo.
 - p. ej. Averías de máquinas en el tiempo, llegadas de aviones a un aeropuerto, pedidos de una empresa, estrellas en el firmamento en cuadrículas del mismo tamaño, etc.
- Supondremos que el proceso que genera estos sucesos se caracteriza por:
 - Es **estable**: i.e. produce a largo plazo un número de sucesos constante por unidad de observación.
 - Los **sucesos aparecen de manera aleatoria** y de forma independiente, i.e. el proceso no tiene memoria:
 - Conocer el número de sucesos en un intervalo no ayuda a predecir el número de sucesos en el siguiente.
- Este proceso es la generalización a un soporte continuo del proceso de Bernoulli.

El proceso de Poisson: La distribución de Poisson

- Dado el proceso anterior la variable aleatoria de Poisson se define como:
 - x = número de sucesos en un intervalo de longitud fija
- La distribución de Poisson aparece como límite de la distribución binomial si suponemos que el número de elementos observados es muy grande pero la probabilidad de observar la característica estudiada en cada elemento es muy pequeña:
 - Dividamos el intervalo de observación t , en n segmentos muy pequeños (así el número de segmentos es muy grande), y observemos el suceso en cada uno de ellos.
 - Si la probabilidad de este suceso, p , es muy pequeña entonces la aparición de dos o más sucesos es despreciable en el segmento.
 - Así el problema se puede plantear como la probabilidad de observar en n elementos si aparece el suceso estudiado o no, i.e. la distribución binomial.
- Es decir la distribución Poisson es un caso límite de esta distribución binomial cuando n tiende a infinito y p tiende a cero, pero de manera que el número medio de sucesos, np , permanezca constante.

El proceso de Poisson: La distribución de Poisson

- Consideremos que el número de accidentes por 100 horas de conducción para un grupo de conductores es λ , y que los accidentes ocurren con el proceso de Poisson: aleatoria e independiente a lo largo del tiempo.
- Así la variable aleatoria x de Poisson será: # accidentes de accidentes en 100 horas de conducción.
- Como hemos dicho antes podemos convertir x en una binomial:
 - Considerando intervalos de tiempo muy pequeños (p. ej. Cada minuto), donde la probabilidad p de ocurrencia de dos accidentes sea despreciable.
 - Así x puede considerarse como una variable binomial en un experimento con $n=100 \times 60$ minutos = 6000 repeticiones, de observar un accidente en un minuto.
 - La probabilidad p de accidente cumple: $E(x) = \lambda = np$, y $p = \lambda/n$.
- Así la distribución Poisson la obtendremos en el caso cuando n tiende a infinito y p tienda a cero, pero de manera que número medio de sucesos, $\lambda = np$, permanezca constante.

El proceso de Poisson: La distribución de Poisson

- En conclusión la distribución de Poisson puede aproximarse por:

- $P(x = r) = \binom{n}{r} \left(\frac{\lambda}{n}\right)^r \left(1 - \frac{\lambda}{n}\right)^{n-r}$, así si tomamos los límites:

- $\lim_{n \rightarrow \infty} P(x = r) = \frac{\lambda^r}{r!} \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-r+1)}{\left(1 - \frac{\lambda}{n}\right)^r n^r} \left(1 - \frac{\lambda}{n}\right)^n \rightarrow (n - \lambda)^r$

- El primer termino del límite : $\frac{n}{(n-\lambda)} \frac{(n-1)}{(n-\lambda)} \dots \frac{(n-r+1)}{(n-\lambda)} = 1$

- El segundo: $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$ r términos

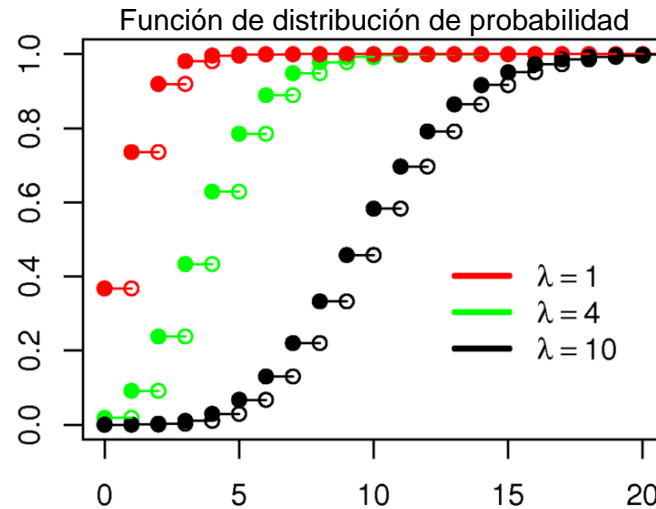
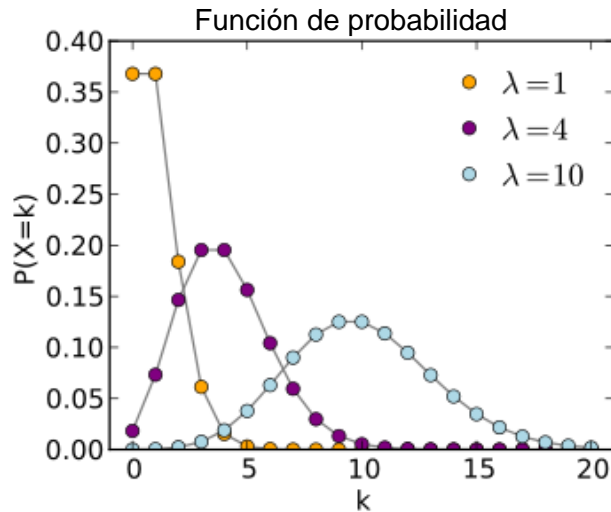
$$\begin{aligned} \binom{n}{r} &= \frac{n!}{r!(n-r)!} \\ &= \frac{n(n-1)\dots(n-r+1)(n-r)!}{r!(n-r)!} \\ &= \frac{n(n-1)\dots(n-r+1)}{r!} \end{aligned}$$

- Por tanto tenemos en el límite $n \rightarrow \infty$, con λ constante:

- $P(x = r) = \frac{\lambda^r}{r!} e^{-\lambda}, r = 0, 1, 2, \dots$ Probabilidad de encontrar r sucesos en un intervalo de longitud fija n que tiene implícito $\lambda=np$, donde p es la probabilidad del suceso observado, (**ver ejemplo 5.4**).

El proceso de Poisson: La distribución de Poisson

- Las medidas características de la distribución de Poisson son:
 - $E[x]=Var[x]=\lambda$. (Ver apéndice 5A).
- Observar que los resultados son consistentes con la aproximación binomial:
 - La varianza de la binomial es npq y cuando $n \rightarrow \infty$ y $p \rightarrow 0$, pero con $np=\lambda=cte$, entonces $q \rightarrow 1$, lo que implica $npq \rightarrow np=\lambda$, que es la varianza de Poisson.



Imágenes extraídas de
https://es.wikipedia.org/wiki/Distribuci%C3%B3n_de_Poisson

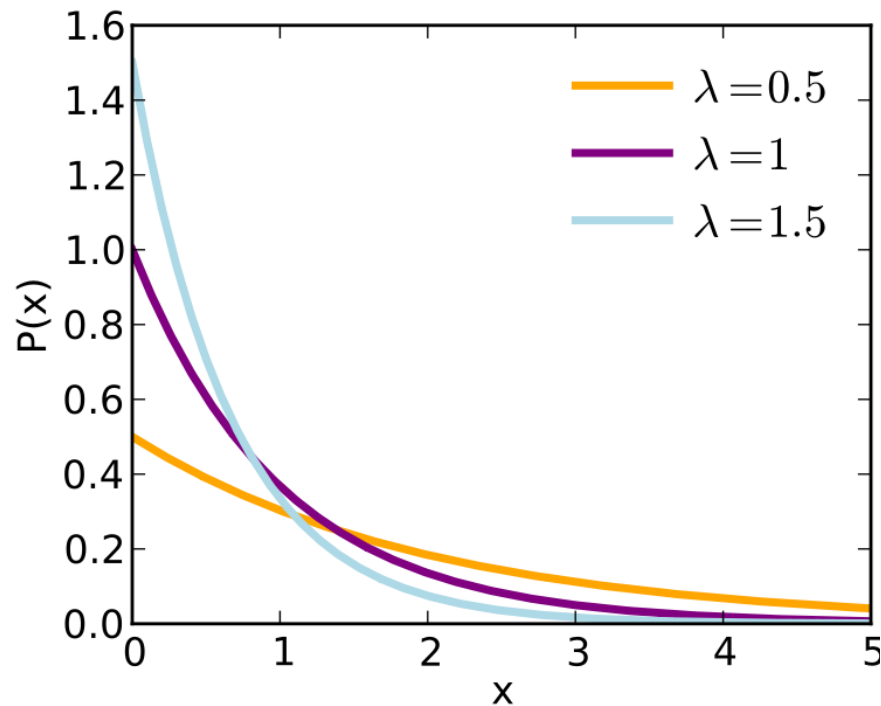
El proceso de Poisson: La distribución de exponencial

- La variable exponencial resulta al considerar en un proceso de Poisson la variable continua:
 - t =tiempo entre la ocurrencia de dos sucesos consecutivos
- Esta variable toma valores en el intervalo $(0, \infty)$.
- Para obtener la función distribución de esta variable primero observemos que la probabilidad de que en un proceso de Poisson de media λt no se produzca ningún suceso en un tiempo t es:
 - $P(t > t_0) = P[\text{cero sucesos en } (0, t_0)] = e^{-\lambda t_0}$, siendo λ la tasa media de sucesos por unidad de tiempo.
 - Observar que $P(t > t_0)$ es la probabilidad de que el tiempo, t , de ocurrencia entre dos sucesos consecutivos sea mayor que un cierto t_0 , i.e. que entre 0 y t_0 no ocurran dos sucesos consecutivos.
 - Así la función de distribución $F(t_0) = P(t \leq t_0) = 1 - e^{-\lambda t_0}$.
 - Cuya función de densidad será: $f(t) = dF(t)/dt = \lambda e^{-\lambda t}$, $\lambda > 0$ y $t > 0$.
 - Las medidas características son: $E[t] = 1/\lambda = DT[t]$, $Var[t] = 1/\lambda^2$
 - Ver **ejemplo 5.5** del libro.

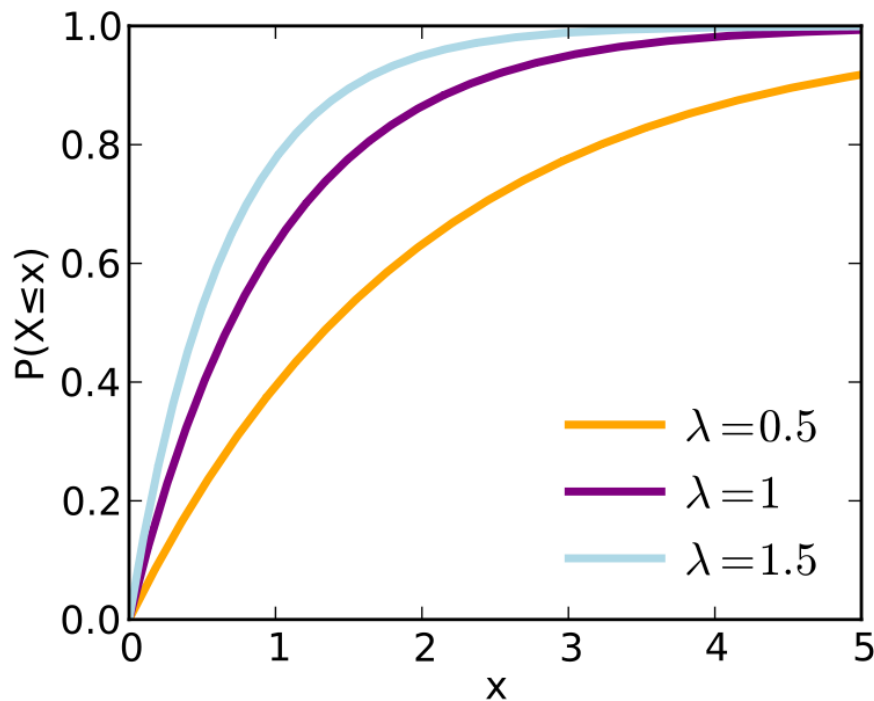
$P(t > t_0)$ se entiende función de supervivencia: la probabilidad de que una variable aleatoria t tome valores superiores a un valor dado t_0

El proceso de Poisson: La distribución de exponencial

Función de probabilidad



Función de distribución de probabilidad



Imágenes extraídas de

https://es.wikipedia.org/wiki/Distribuci%C3%B3n_exponencial

Distribuciones de duraciones de vida

- La distribución exponencial es el ejemplo más simple para distribuciones de variables aleatorias continuas que pueden tomar cualquier valor positivo no acotado.
- Se utilizan para modelar duración (vida de personas, animales o componentes físicos, duraciones de huelgas, periodos de desempleos, duración de discursos políticos, etc.), o el tamaño (rentas de familias, tamaño de yacimientos, etc.).
- En el caso particular del modelado de duración, estas son las duraciones de vida de ciertos elementos, que es en el caso que nos centraremos a continuación.
- Ver ejemplo 5-5 del libro.
- Estudiaremos ahora que cuando medimos la duración de una vida de un elemento determinado, esta la podemos caracterizar por la función que proporciona la probabilidad de muerte en cada instante (denominada **tasa de fallos**, definiendo la función densidad de la variable) para los elementos que han sobrevivido hasta dicho instante.
- Así supongamos que $f(t)$ es la función de densidad de una variable continua positiva en $(0, \infty)$, que representa por ejemplo la **duración de vida** de ciertos elementos (así como la del **ejemplo 5-5**).

Distribuciones de duraciones de vida

- Así podemos calcular la **probabilidad de muerte o fallo** en el intervalo $(t_0, t_0 + \Delta t)$ para los elementos que ya han vivido hasta t_0 , mediante la siguiente probabilidad condicionada:
 - $P(t_0 < t \leq t_0 + \Delta t \mid t > t_0) = P(t_0 < t \leq t_0 + \Delta t, t > t_0) / P(t > t_0) = P(t_0 < t \leq t_0 + \Delta t) / P(t > t_0)$. A la función $P(t > t_0)$ se le suele llamar función de **fiabilidad o supervivencia** y se define como $P(t > t_0) = \int_{t_0}^{\infty} f(t) dt = 1 - F(t_0)$.
 - El numerador de la probabilidad condicional lo tomamos como $P(t_0 < t \leq t_0 + \Delta t, t > t_0) = P(t_0 < t \leq t_0 + \Delta t)$ ya que $P(t_0 < t \leq t_0 + \Delta t)$ coincide probabilidad conjunta de los dos sucesos:
 - $t > t_0$ y $t_0 < t \leq t_0 + \Delta t$, i.e. los intervalos solapan (el suceso $t_0 < t \leq t_0 + \Delta t$, esta incluido en el suceso $t > t_0$).
- Si definimos $F(t_0)$ como la función distribución de la variable t_0 , recordemos que la función de distribución es $F(t_0) = P(t \leq t_0)$, así podemos aproximar:
 - $P(t_0 < t \leq t_0 + \Delta t \mid t > t_0) \approx f(t_0)\Delta t / 1 - F(t_0)$, ya que la probabilidad en un intervalo viene dado por el área que encierra la densidad de probabilidad, $f(x)$. i.e. un rectángulo de base Δt y de altura $f(t_0)$.

Distribuciones de duraciones de vida

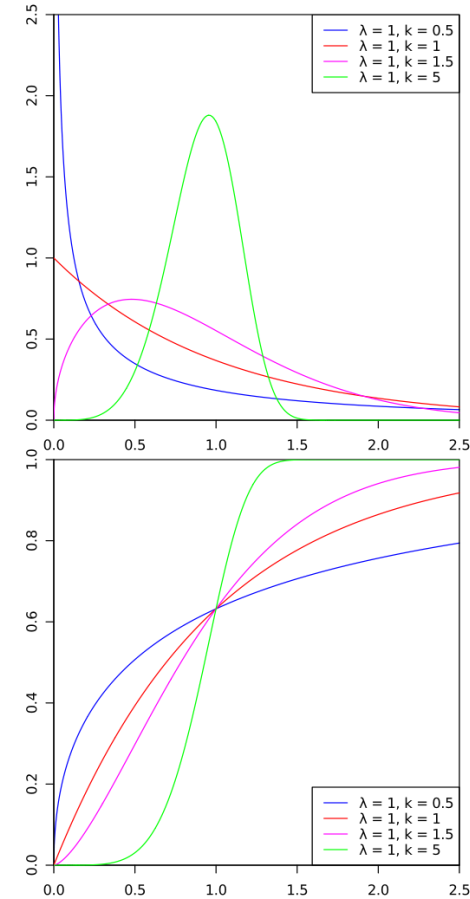
- Así podemos definir $\lambda(t)$ como la **tasa de fallo** instantánea en el límite $\Delta t \rightarrow 0$:
- $\lambda(t) = f(t) / 1 - F(t)$, la tasa de fallos se define como $\lim_{\Delta t \rightarrow 0} P(t_0 < t \leq t_0 + \Delta t | t > t_0) / \Delta t$.
- Esta cantidad representa la **probabilidad de muerte en cada instante para los elementos que han sobrevivido hasta dicho instante (por tanto depende de t)**.
- Para obtener la función densidad $f(t)$ solo en función de la tasa de fallo, $\lambda(t)$, sabiendo que $F(0) = 0$, integramos (definimos $\Lambda(t)$ es la función de **tasa de fallo acumulada**):
 - $\Lambda(t) = \int_0^t \lambda(x) dx = \int_0^t \frac{f(x)}{1 - F(x)} dx = -\ln[1 - F(x)]_0^t = -\ln[1 - F(t)]$.
 - Así $1 - F(t) = \exp\{-\Lambda(t)\}$ y $F(t) = 1 - \exp\{-\Lambda(t)\}$ y derivando
 - $f(t) = \lambda(t) \exp\{-\Lambda(t)\}$, que es la forma habitual de las distribuciones continuas para variables positivas para el tiempo de vida media de un cierto suceso.
- Esta es la función densidad de la variable aleatoria que regula la tasa de fallo. En este caso la probabilidad de morir en cualquier intervalo depende de la vida anterior, por eso λ depende del tiempo.

Distribuciones de duraciones de vida

- La distribución exponencial pura se caracteriza por una tasa de fallo constante: la probabilidad de morir en cualquier intervalo no depende de la vida anterior, i.e. $f(t) = \lambda e^{-\lambda t}$, $\lambda > 0$ y $t > 0$, λ es constante.
- Darse cuenta que para la distribución exponencial $\Lambda(t) = \int_0^t \lambda dt = \lambda t + c$ y tomando arbitrariamente $c=0$, tenemos que $\Lambda(t) = \lambda t$, así para el caso particular de la distribución exponencial tenemos, $f(t) = \lambda e^{-\lambda t} = \lambda \exp\{-\Lambda(t)\}$.
- Así la **distribución exponencial** se puede considerar como una distribución de duración de vida para una **tasa de fallo λ constante**: $f(t) = \lambda \exp\{-\Lambda(t)\} = \lambda e^{-\lambda t}$
- Por tanto la distribución exponencial es adecuada para describir aparición de muertes al azar, no debidas a desgaste o deterioro.
- No obstante podemos suponer una **tasa de fallo no constante** del tipo: $\lambda(t) = ht^{c-1}$
 - Tenemos una tasa de fallo que aumentará con el tiempo si $c > 1$.
 - Será constante (distribución exponencial con $\lambda(t) = h$) si $c = 1$ (distribución exponencial).
 - Y disminuirá si $c < 1$.
 - Siendo la función densidad dada por: $f(t) = ht^{c-1} \exp\{(-h/c)t^c\}$.
 - Esta es **distribución de Weibull** (ver ejemplos en el libro).

Distribuciones de duraciones de vida: distribución de Weibull

- Normalmente se pueden definir unos parámetros de forma y escala para la distribución de Weibull en la tasa acumulada, siendo el parámetro de **forma** k y parámetro de **escala** λ . Así se puede definir la función de tasas acumulada $\Lambda(x) = \left(\frac{x}{\lambda}\right)^k$ y por tanto la función de distribución queda para $x \geq 0$:
- $$f(x; \lambda, k) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^k\right\},$$
- recordar que la tasa de fallos es la derivada de la tasa de fallos acumulada, i.e. $d\Lambda(x)/dx$.



Imágenes extraídas de

https://es.wikipedia.org/wiki/Distribuci%C3%B3n_de_Weibull