Postgrado en Modelado y Simulación

DISTRIBUCIONES DE PROBABILIDAD

PROF. MAGDIEL ABLAN BORTONE

PROF. PAOLO RAMONI PERAZZI

CESIMO
Centro de Simulación y Modelos

FACULTAD DE INGENIERÍA

UNIVERSIDAD DE LOS ANDES

Modelos probabilísticos

- Modelos probabilísticos discretos
  - Distribución binomial
  - Distribución geométrica
  - Distribución binomial negativa
  - Distribución de Poisson
- Modelos probabilísticos continuos
  - Distribución uniforme
  - Distribución exponencial
  - Otras distribuciones (Pareto, Weibull, Gamma)
  - Distribución normal

VVAA discretas

- Son las llamadas distribuciones de conteo
- Se usan para modelizar la frecuencia de los eventos y el tiempo de espera en el caso discreto
- Muchas de las distribuciones más importantes se definen tomando en cuenta las pruebas o ensayos de Bernoulli

Ensayo de Bernoulli

- Experimento aleatorio con sólo dos resultados posibles, tradicionalmente denominados éxito y fracaso

- Una VA X es Bernoulli si

$$p(0) = P(X = 0) = 1 - p$$

$$p(1) = P(X = 1) = p$$

- Donde p es la probabilidad de éxito y $0 < p < 1$

- $E[X] = p$ y $Var[X] = p(1\text{-}p)$

Distribución binomial

Postgrado en Modelado y Simulación

- Sea X el número de éxitos en n ensayos de Bernoulli, independientes e idénticamente distribuidos, con probabilidad de éxito p, se dice que $X \sim bin(n,p)$ tal que $P(X = i) = p(i)$ es:

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0,1,\dots,n$$

- Donde $\binom{n}{i} = \frac{n!}{(n-i)!i!}$
- $E[X] = np$ y $Var[X] = np(1-p)$

Distribución binomial

- Ejemplo: entrecruzar flores rojas y blancas produce 25% de flores rojas
- Cruzamos cinco pares de flores rojas y blancos y producen cinco descendientes
- ¿Cuál es la p que no se obtengan hijos rojos?

$$P(X=0)=\frac{5!}{(5-0)!\,0!}\,p^{0}(1-p)^{5}$$

$$=1(0.25)^{0}(0.75)^{0}=0.237$$

```
>dbinom(0, 5, 0.25)
```

Distribución binomial

- ¿Cuál es la probabilidad que se obtengan menos de dos hijos rojos?

$$P(X < 2) = P(X = 0) + P(X = 1)$$

$$\frac{5!}{(5-0)!\,0!}\,0.25^{0}(1-0.25)^{5}+\frac{5!}{(5-1)!\,1!}\,0.25^{1}(1-0.25)^{4}$$

$$= 0.237+0.395 = 0.632$$

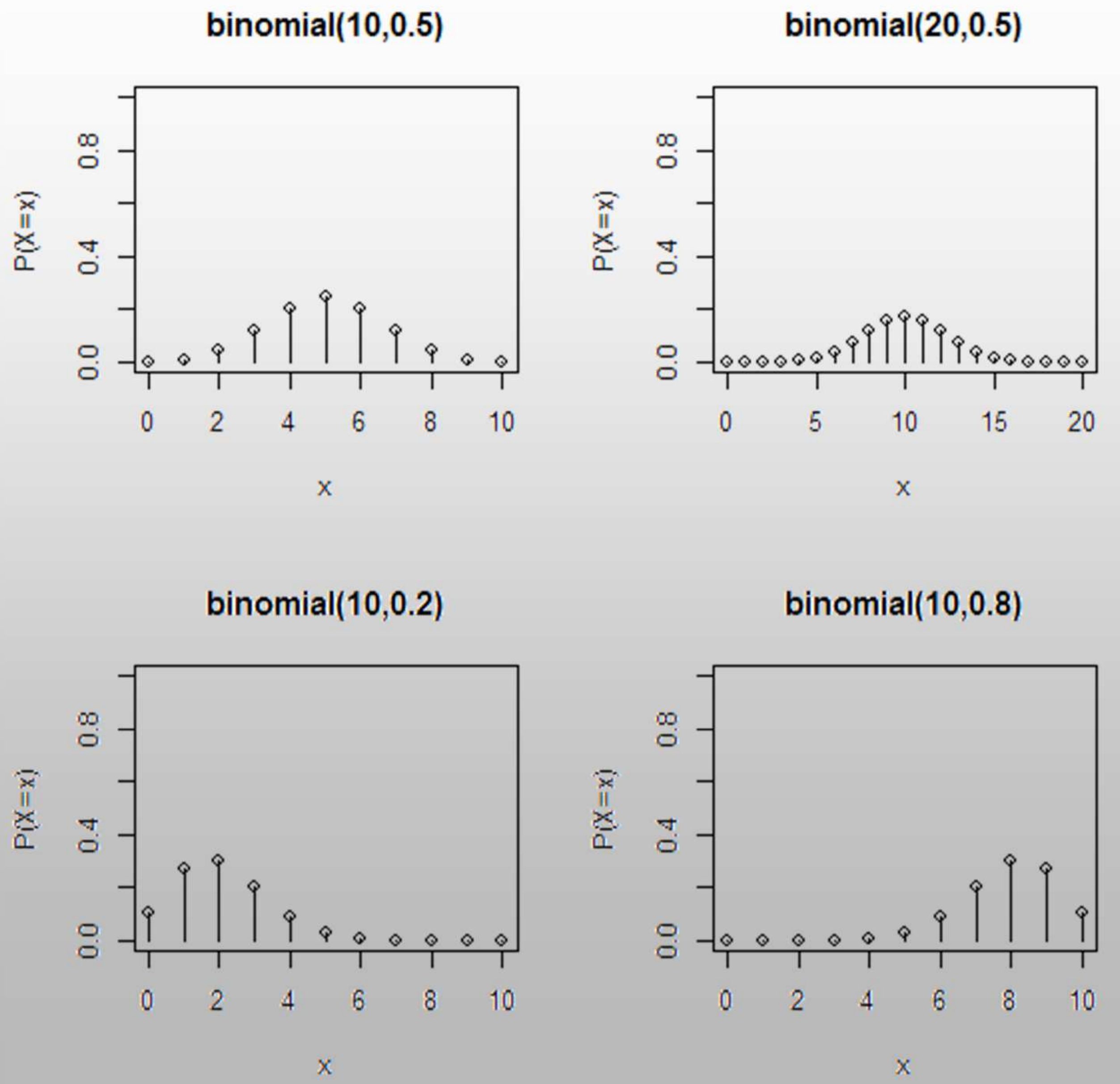
```
>pbinom(1, 5, 0.25)
```

Distribución binomial

Postgrado en Modelado y Simulación

```
> par(mfrow=c(2,2))
> plot(0:10,dbinom(0:10,10,0.5),type="h",xlab="x",yla
      b="P(X=x)",ylim=c(0,1),main="binomial(10,0.5)")
> points(0:10,dbinom(0:10,10,0.5),pch=1)

> plot(0:20,dbinom(0:20,20,0.5),type="h",xlab="x",yla
      b="P(X=x)",ylim=c(0,1),main="binomial(20,0.5)")
> points(0:20,dbinom(0:20,20,0.5),pch=1)

> plot(0:10,dbinom(0:10,10,0.2),type="h",xlab="x",yla
      b="P(X=x)",ylim=c(0,1),main="binomial(10,0.2)")
> points(0:10,dbinom(0:10,10,0.2),pch=1)

> plot(0:10,dbinom(0:10,10,0.8),type="h",xlab="x",yla
      b="P(X=x)",ylim=c(0,1),main="binomial(10,0.8)")
> points(0:10,dbinom(0:10,10,0.8),pch=1)
```

Distribución binomial



Distribución binomial

- Otro ejemplo: se lanzan cuatro monedas justas. Los resultados se presumen independientes. ¿Cuál es la p de obtener dos caras y dos sellos?
- X es el número de caras (éxitos)
- Parámetros n = 4 y p = 1/2
- ¿$P(X=2)$?

```
>dbinom(2,4,0.5)
[1] 0.375
```

Distribución geométrica

- Considere una secuencia de ensayos de Bernoulli con probabilidad de éxito p. se define la VA X como el número de fallas hasta observar el primer éxito. entonces:

$$P(X = x) = p(1 - p)^x, x = 0, 1, 2, \dots$$

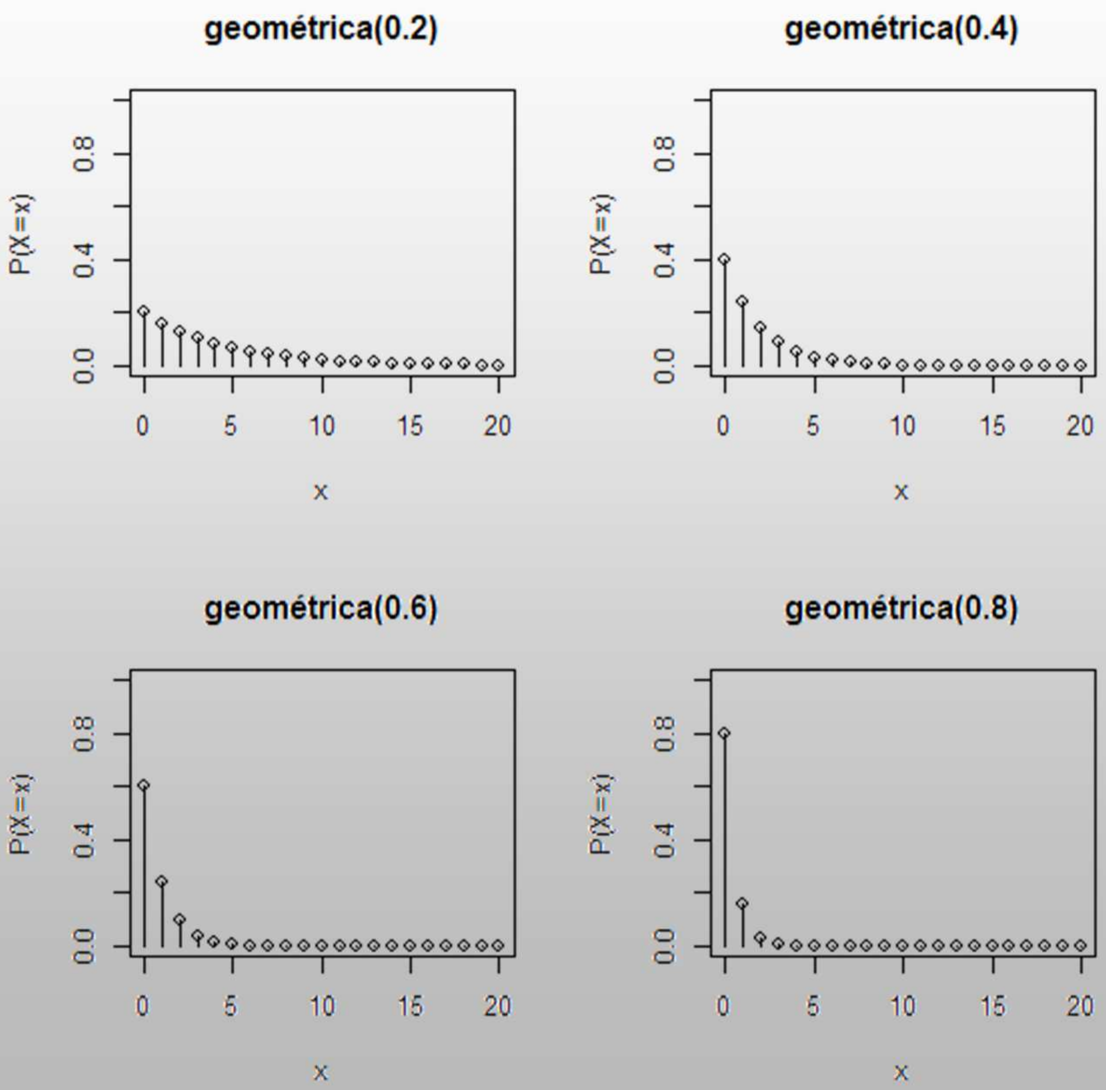
- Y se denota como $X \sim geom(p)$
- $E[X] = (1\text{-}p)/p$ y $Var[X] = (1\text{-}p)/p2$
- Ojo: algunas veces se define como el número de ensayos hasta obtener el primer éxito, Así:

$$P(X = x) = p(1 - p)^{x-1}, x = 0, 1, 2, \dots$$

Distribución geométrica

```
> opar<-par(mfrow=c(2,2))
> plot(0:20,dgeom(0:20,0.2),type="h",xlab="x",ylab="P
  (X=x)",ylim=c(0,1),main="geométrica(0.2)")
> points(0:20,dgeom(0:20,0.2),pch=1)

> plot(0:20,dgeom(0:20,0.4),type="h",xlab="x",ylab="P
  (X=x)",ylim=c(0,1),main="geométrica(0.4)")
> points(0:20,dgeom(0:20,0.4),pch=1)

> plot(0:20,dgeom(0:20,0.6),type="h",xlab="x",ylab="P
  (X=x)",ylim=c(0,1),main="geométrica(0.6)")
> points(0:20,dgeom(0:20,0.6),pch=1)

> plot(0:20,dgeom(0:20,0.8),type="h",xlab="x",ylab="P
  (X=x)",ylim=c(0,1),main="geométrica(0.8)")
> points(0:20,dgeom(0:20,0.8),pch=1)
> par(opar)
```

Distribución geométrica



Distribución geométrica

- Ejemplo, está tratando de encender una parrilla con fósforos en un día con mucho viento. Cada fósforo tiene una p de encender la parrilla de 0.1 pero sólo hay 4. ¿Cuál es la p que encienda la parrilla antes de quedarse sin fósforos?
- Suponga que tiene una cantidad ilimitada de fósforos y sea X el número de intentos fallidos antes de encender la parrilla.
- Entonces $X \sim geom(p=0.1)$ y se pide calcular $P(X \leq 3) = pgeom(3,0.1) = 0.3439$

Distribución geométrica

- Ahora suponga que si se usan dos fósforos al mismo tiempo, la p de encendido es 0.3
- ¿Es una buena idea usar los dos fósforos al mismo tiempo?
- Sea Y la VA que representa el número de intentos fallidos hasta encender la parrilla usando dos fósforos a la vez
- Si $P\ (Y \leq 1)$ es más grande que $P\ (X \leq 3)$ valdrá la pena usar dos fósforos en vez de uno

Distribución binomial negativa

- Sea Z el número de fallas hasta obtener el r-ésimo éxito en una secuencia de ensayos de Bernoulli independientes
- Se dice entonces que $Z \sim negbin(r,p)$ y su función de masa de probabilidad está dada por:

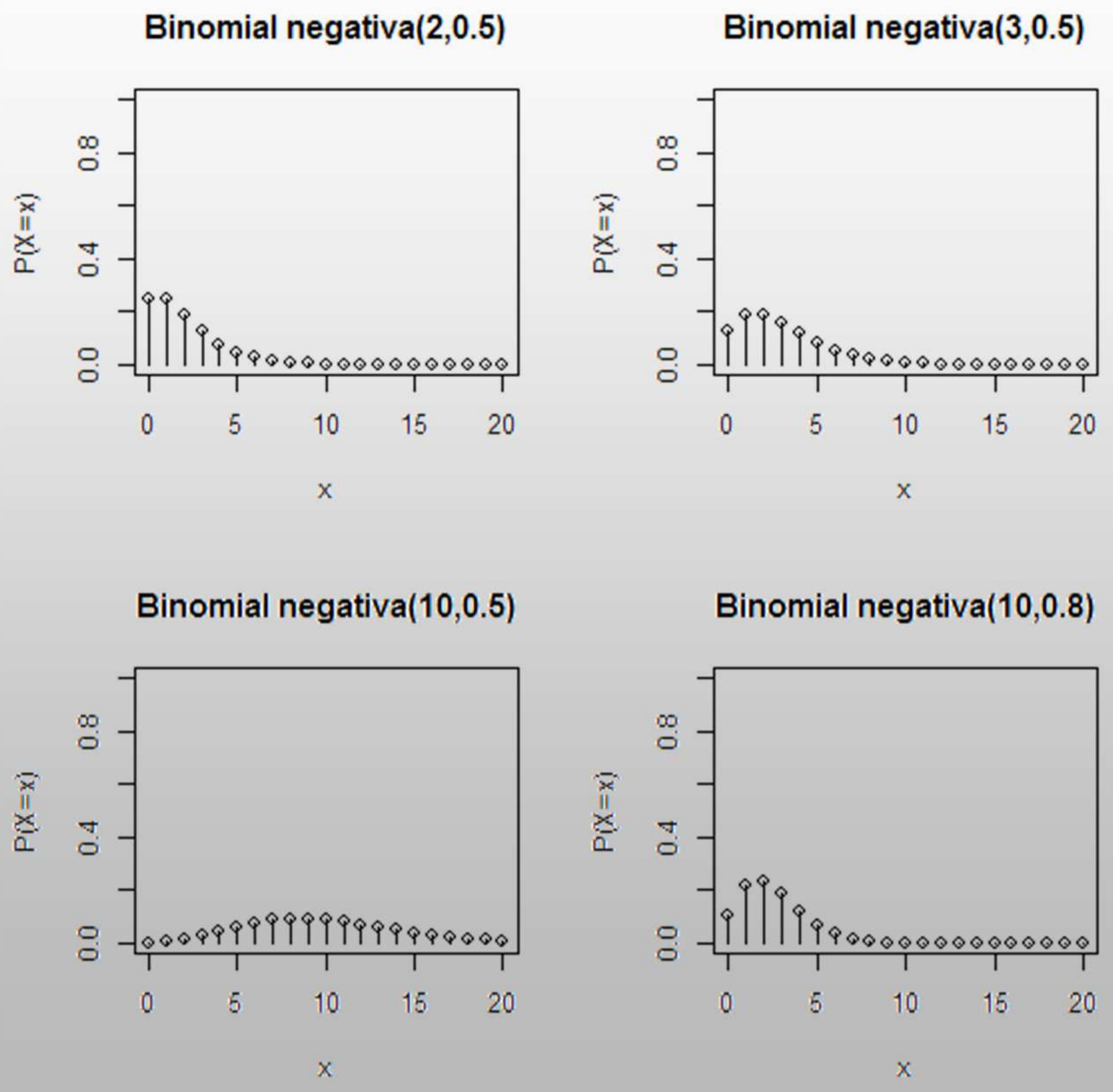
$$P(Z=z)=\binom{r+z-1}{r-1}p^r(1-p)^z \quad para\ z=0,1\ldots$$

- $E[Z] = r(1-p)/p$ y $Var[Z] = r(1-p)/p^2$
- Igual que la geométrica, a veces se define como el número de ensayos y cambia la expresión de probabilidad

Distribución binomial negativa


```
> par(mfrow=c(2,2))
> plot(0:20,dnbinom(0:20,2,0.5),type="h",xlab="x",ylab="P(X
  =x)",ylim=c(0,1),main="Binomial negativa(2,0.5)")
> points(0:20,dnbinom(0:20,2,0.5),pch=1)

> plot(0:20,dnbinom(0:20,3,0.5),type="h",xlab="x",ylab="P(X
  =x)",ylim=c(0,1),main="Binomial negativa(3,0.5)")
> points(0:20,dnbinom(0:20,3,0.5),pch=1)

> plot(0:20,dnbinom(0:20,10,0.5),type="h",xlab="x",ylab="P(
  X=x)",ylim=c(0,1),main="Binomial negativa(10,0.5)")
> points(0:20,dnbinom(0:20,10,0.5),pch=1)

> plot(0:20,dnbinom(0:20,10,0.8),type="h",xlab="x",ylab="P(
  X=x)",ylim=c(0,1),main="Binomial negativa(10,0.8)")
> points(0:20,dnbinom(0:20,10,0.8),pch=1)
```

Distribución binomial negativa



Distribución binomial negativa

- Un fabricante prueba la calidad de su producto seleccionando al azar una muestra de tamaño 100 de cada lote
- Si hay más de dos artículos defectuosos, la producción se detiene para solucionar el problema
- Cada artículo es defectuoso con independencia de los otros con probabilidad $p = 0.01$
- En la práctica, los artículos en la muestra se prueban secuencialmente.

Distribución binomial negativa

- El procedimiento se para cuando se consiguen tres artículos defectuosos.
- Sea Z el número de artículos observados antes de conseguir tres artículos defectuosos. Entonces $Z \sim negbin(3,0.01)$
- ¿Cuál es la p de que se pare la producción?
- $P(parar\ la\ producción) = p(Z + 3 \leq 100) = p(z \leq 97)$

```
>pnbinom(97,3,0.01)
[1] 0.0793732
```

Distribución de Poisson

- X tiene distribución de Poisson con parámetro λ, $X \sim pois(\lambda)$ si:

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \qquad x = 0, 1, \ldots$$

$$E[x] = \lambda \text{ y } Var[x] = \lambda$$

- La distribución de Poisson se usa como modelo para eventos raros y eventos que ocurren al azar a una tasa constante en el tiempo o el espacio.

Distribución de Poisson

Postgrado en Modelado y Simulación

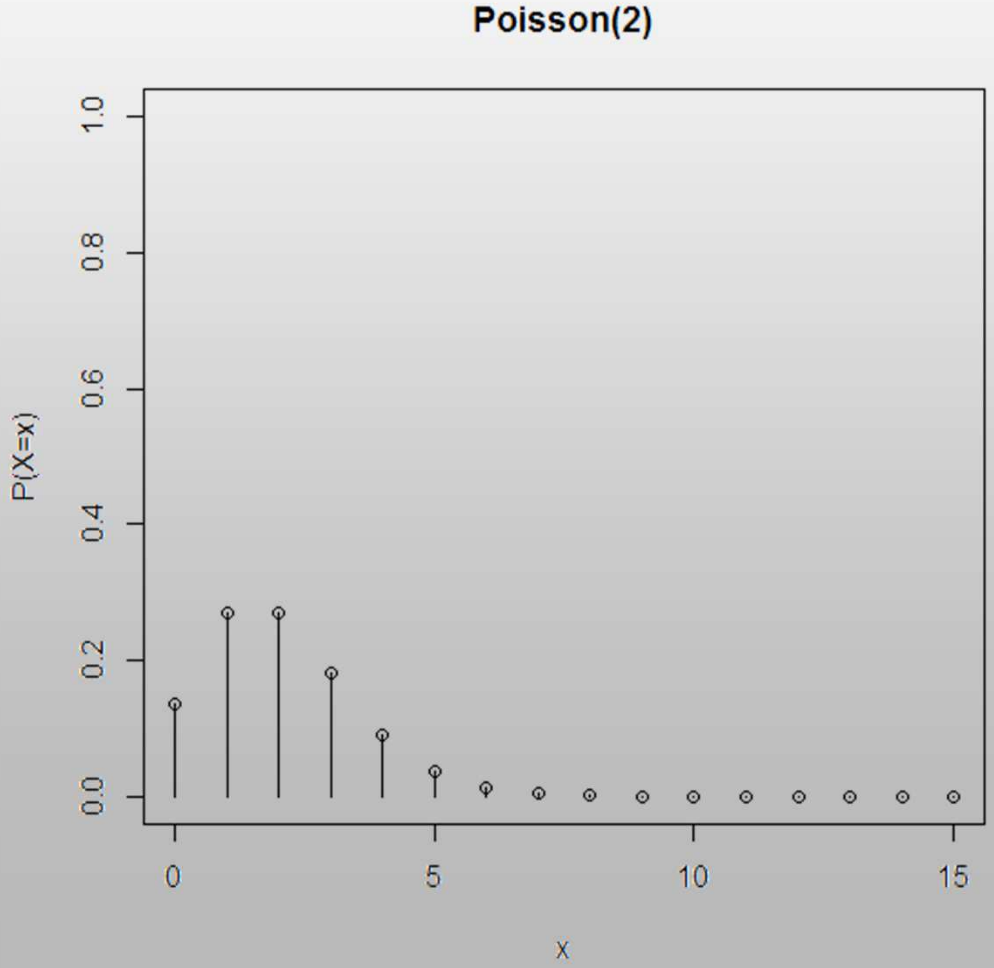
- Los ejemplos clásicos incluyen al número de accidentes en un año, el número de llamadas que llegan a una central telefónica en una hora dada, el número de carros que llegan a una intersección dada, entre otros.

- Ejemplo: el n medio de accidentes ocurridos en un planta petrolera es 2 accidentes en 2 meses

1. ¿Qué modelo sigue la variable n de accidentes ocurridos en la planta por 2 meses?.
2. ¿Cuál es la p que haya más de 2 accidentes en 2 meses?
3. ¿Cuál que haya entre 2 y 8 inclusive, en 2 meses?
4. ¿Cuál que haya más de 2 en 1 mes?

Distribución de Poisson

1. Es razonable pensar que sigue una distribución de Poisson de parámetro $\lambda = E[X] = 2$




```
plot(0:15,dpois(0:15,2),type="h",
   xlab="x",ylab="P(X=x)",ylim=c(0,1),
   main="Poisson(2)")
points(0:15,dpois(0:15,2),pch=1)
```

Distribución de Poisson

2. $P(X > 2) = 1 - P(X \leq 2) = 1 - 0.6767 = 0.3233$

```
> 1-ppois(2,2)
```

3. $P(2 \leq X \leq 8) = P(X \leq 8) - P(X \leq 1) = 0.999 - 0.406 = 0.5938$

```
> ppois(8,2)-ppois(1,2)
```

4. $Y \sim pois(1)$

   $P(Y > 2) = 1 – P(Y \leq 2) = 1 – 0.9197 = 0.0803$

```
>1-ppois(2,1)
```

Distribución de Poisson

Postgrado en Modelado y Simulación

- El n de errores tipográficos en una página dada de estas láminas se distribuye como una Poisson de parámetro λ=1. ¿Cuál es la probabilidad de conseguir uno o más errores? Es decir $P(X \geq 1)$
- Estas expresiones en R son equivalentes y arrojan el valor correcto `[1] 0.633`

```
>ppois(0,1,lower.tail=F)

>1- ppois(0,1)

>dpois(0,1)
```

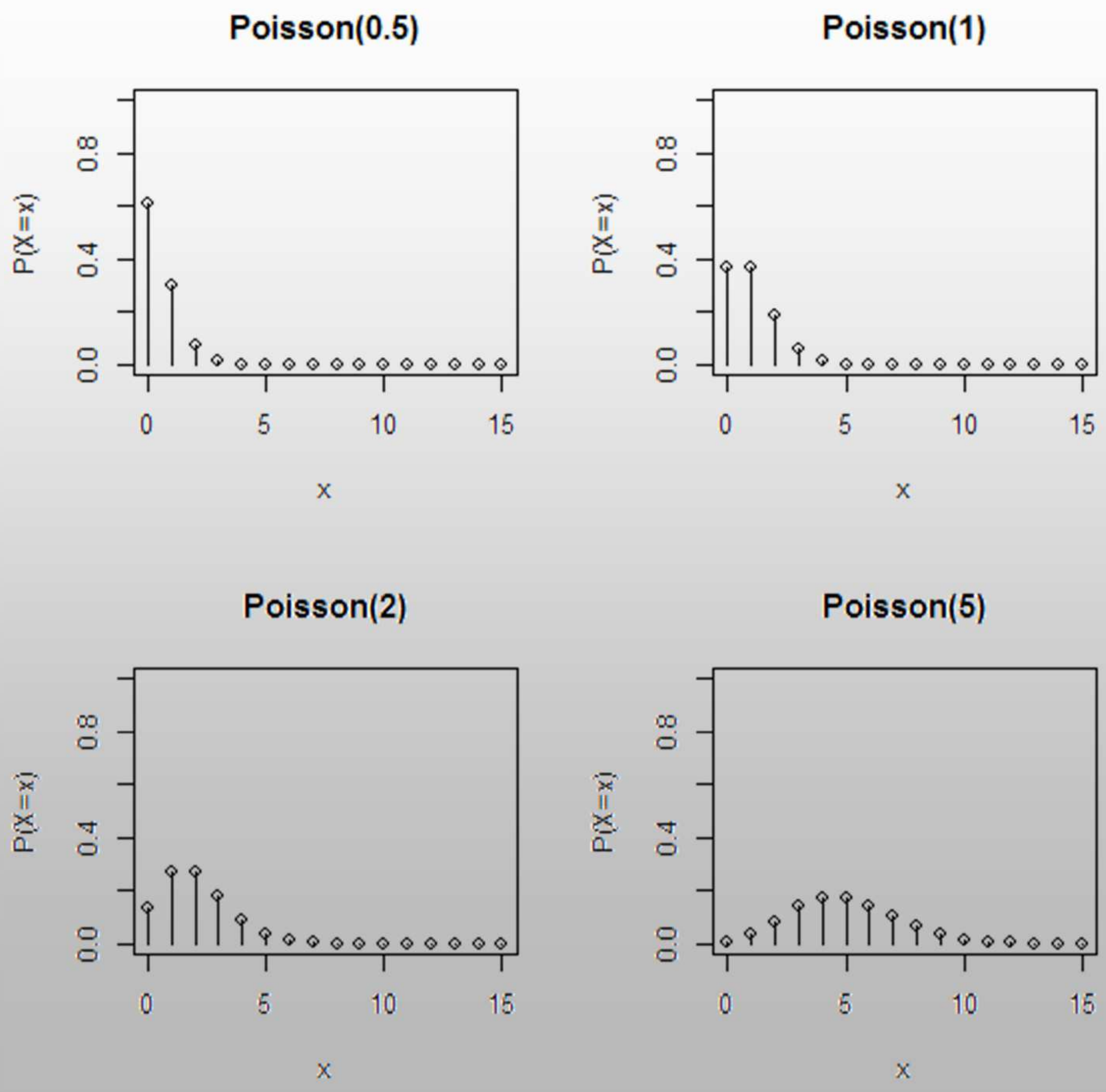
Distribución de Poisson


```
> par(mfrow=c(2,2))

> plot(0:15,dpois(0:15,0.5),type="h",xlab="x",ylab="P
   (X=x)",ylim=c(0,1),main="Poisson(0.5)")
> points(0:15,dpois(0:15,0.5),pch=1)

> plot(0:15,dpois(0:15,1),type="h",xlab="x",ylab="P(X
    =x)",ylim=c(0,1),main="Poisson(1)")
> points(0:15,dpois(0:15,1),pch=1)

> plot(0:15,dpois(0:15,2),type="h",xlab="x",ylab="P(X
    =x)",ylim=c(0,1),main="Poisson(2)")
> points(0:15,dpois(0:15,2),pch=1)

> plot(0:15,dpois(0:15,5),type="h",xlab="x",ylab="P(X
    =x)",ylim=c(0,1),main="Poisson(5)")
> points(0:15,dpois(0:15,5),pch=1)
```

Distribución de Poisson

```
> par(mfrow=c(2,2))

> plot(0:15,dpois(0:15,0.5),type="h",xlab="x",ylab="P
    (X=x)",ylim=c(0,1),main="Poisson(0.5)")
> points(0:15,dpois(0:15,0.5),pch=1)

> plot(0:15,dpois(0:15,1),type="h",xlab="x",ylab="P(X
    =x)",ylim=c(0,1),main="Poisson(1)")
> points(0:15,dpois(0:15,1),pch=1)

> plot(0:15,dpois(0:15,2),type="h",xlab="x",ylab="P(X
    =x)",ylim=c(0,1),main="Poisson(2)")
> points(0:15,dpois(0:15,2),pch=1)

> plot(0:15,dpois(0:15,5),type="h",xlab="x",ylab="P(X
    =x)",ylim=c(0,1),main="Poisson(5)")
> points(0:15,dpois(0:15,5),pch=1)
```

Distribución de Poisson



Distribución Beta

- La distribución Beta es útil cuando lidiamos con proporciones o porcentajes. Veamos un ejemplo:

Promedio de bateo: la proporción entre la cantidad de hits y el número de turnos al bate [0,1], normalmente entre 0,216 y 0,360, con 0,266 como un buen promedio y 0,300 excelente.

- Supongamos: queremos predecir el promedio de un jugador. ¿Podemos usar su promedio de bateo?
- Este problema puede ser representado como una distribución binomial (una serie de aciertos y fallos), y la mejor manera de representar estas expectativas previas es con la distribución beta

Distribución Beta

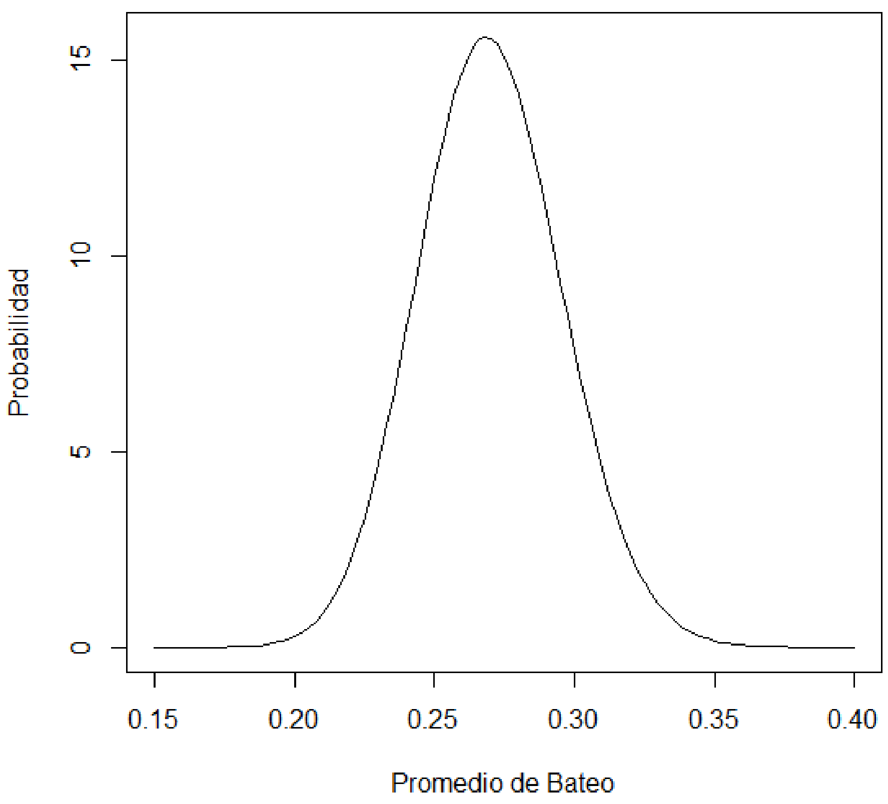
Es decir, antes de ver al jugador lanzar su primer swing, lo que esperamos como su bateo promedio aproximado

- El dominio de la distribución beta [0, 1], igual que las probabilidades.
- Esperamos que el promedio de bateo del jugador sea aproximadamente 0,27, pero razonablemente está entre 0,21 y 0,35
- Esto puede representarse como una distribución beta con parámetros $\alpha=81$ y $\beta=219$, pues:
  - o La esperanza es $\alpha/(\alpha+\beta) = 81/(81+219) = 0.27$
  - o Los límites están entre 0,21 y 0,35…

Distribución Beta


```
curve(dbeta(x, 81, 219), xlim= c(0.15,0.4),
   xlab = "promedio de bateo",
   ylab ="probabilidad")
```

- La distribución Beta está representando una distribución de *probabilidades*
- Pero ¿Por qué es tan apropiada?



Distribución Beta

Imaginemos que el jugador batea un único hit. Ahora su registro para la temporada es un hit/un turno al bate.

- Actualizamos nuestras probabilidades y movemos toda la curva un poquito para incluir la nueva información, la nueva distribución beta es:

$$Beta(\alpha_0+hits,\ \beta_0+fallos)$$

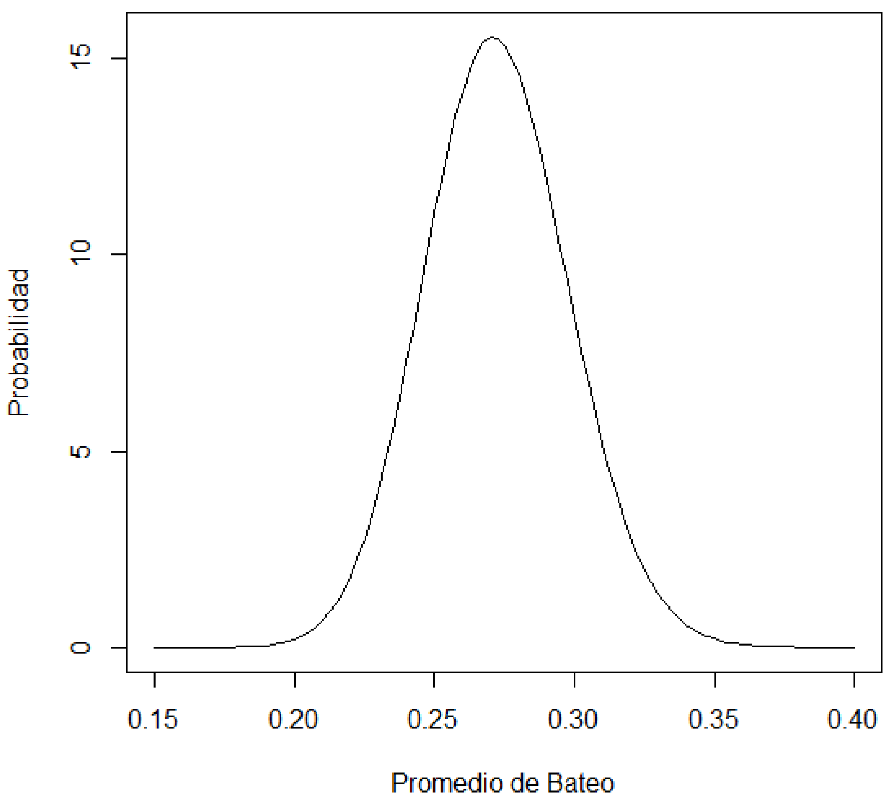
- Donde α_0 y β_0 son los parámetros con los que empezamos (81 y 219)
- En este caso: $beta(81+1,\ 219)$

Distribución Beta

Postgrado en Modelado y Simulación

```
curve(dbeta(x, 82, 219), xlim= c(0.15,0.4),
    xlab = "promedio de bateo",
    ylab ="probabilidad")
```

- Vemos que el cambio es imperceptible
- Pero mientras más juegue, la curva se seguirá acomodando a la nueva evidencia, se irá estrechando en la base en la medida que tengamos más evidencia

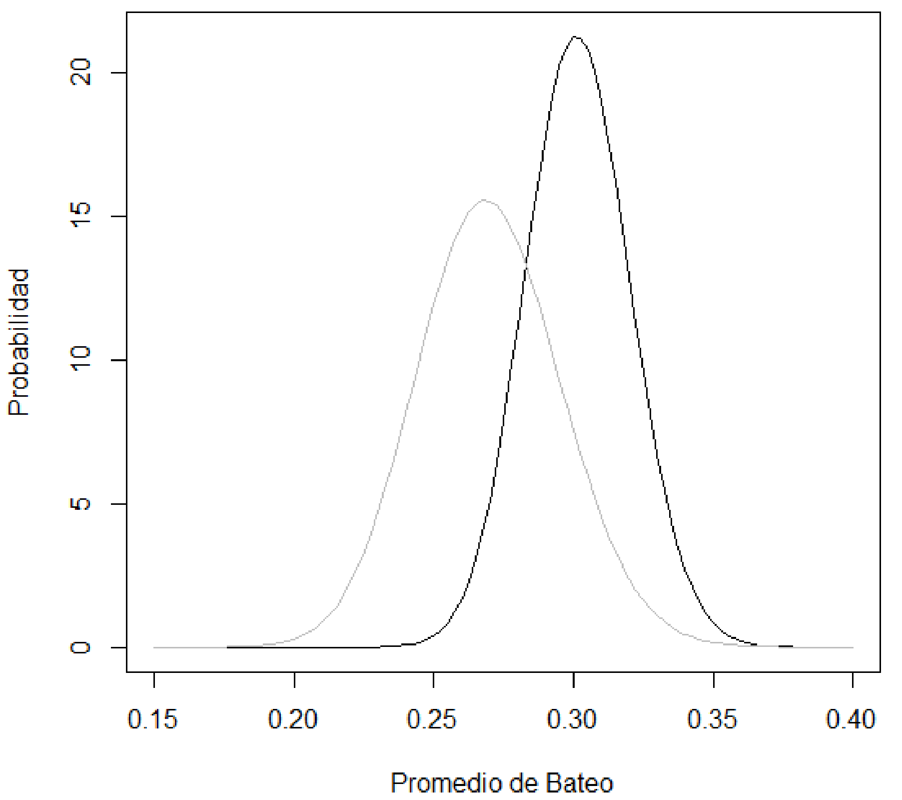


Distribución Beta

```
curve(dbeta(x, 181, 419), xlim= c(0.15,0.4),
    xlab = "promedio de bateo",
    ylab ="probabilidad")
```

Supongamos que a lo largo de la temporada él tuvo 300 turnos al bate y tuvo 100 hits.

- La nueva distribución Beta será...



Distribución Beta

- Una de las salidas más interesantes de esta formula es la esperanza de la distribución Beta resultante, la cuál es el nuevo estimado.
- Recordemos que la esperanza de la distribución Beta es $\alpha/\alpha+\beta$
- Así, luego de 100 hits en 300 turnos al bate, la esperanza es 181/181+419 = 0.303
- Note que es mayor que el estimado inicial de = 0.270).

Distribución uniforme

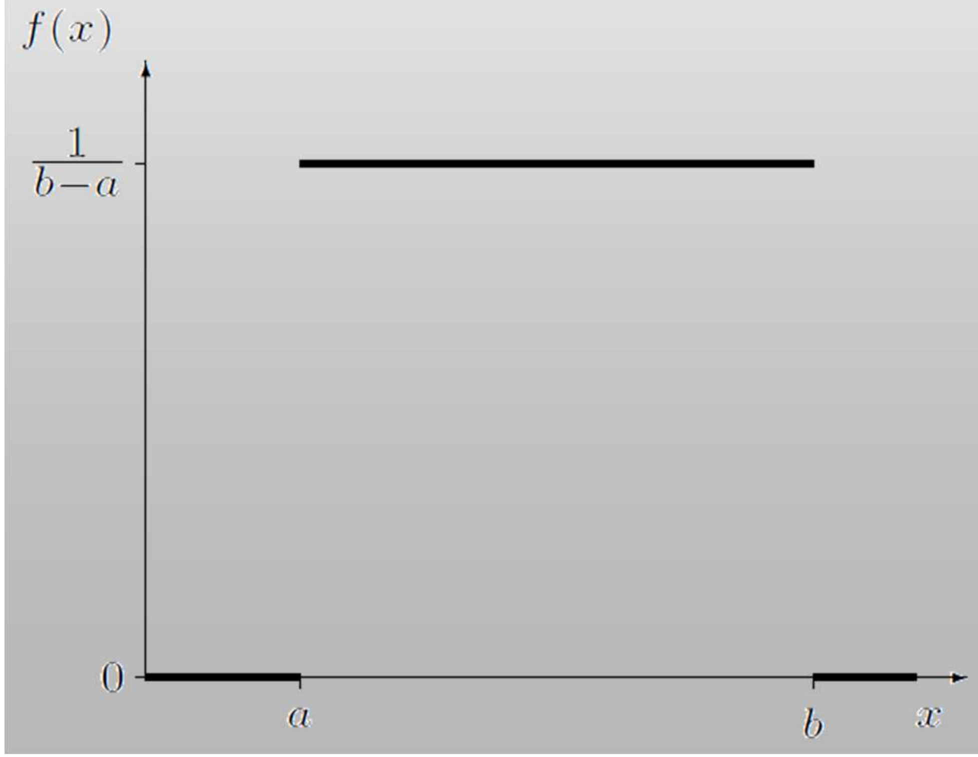
Postgrado en Modelado y Simulación

- La distribución uniforme se usa como un primer modelo para una cantidad que varia aleatoriamente entre a y b pero de la cual no se conoce nada más.
- Además, es esencial en la simulación de todas las otras variables aleatorias

Distribución uniforme

- Cuando la p que X se encuentre en un subintervalo dado [a,b] depende sólo de la longitud del intervalo y no de su localización.

$$X \sim \mathrm{unif}(a, b)$$

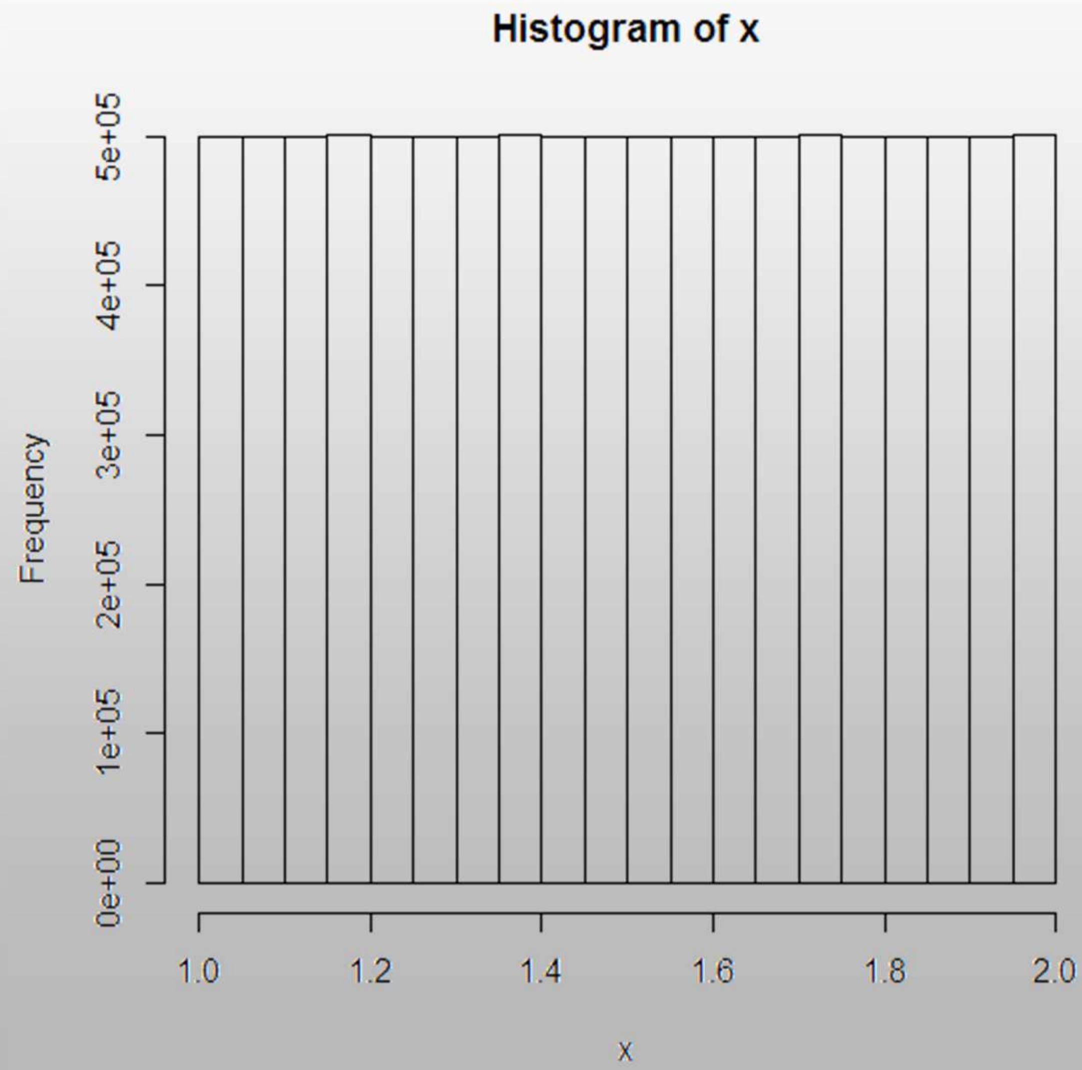


$$\mu = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a)^2}{12}$$

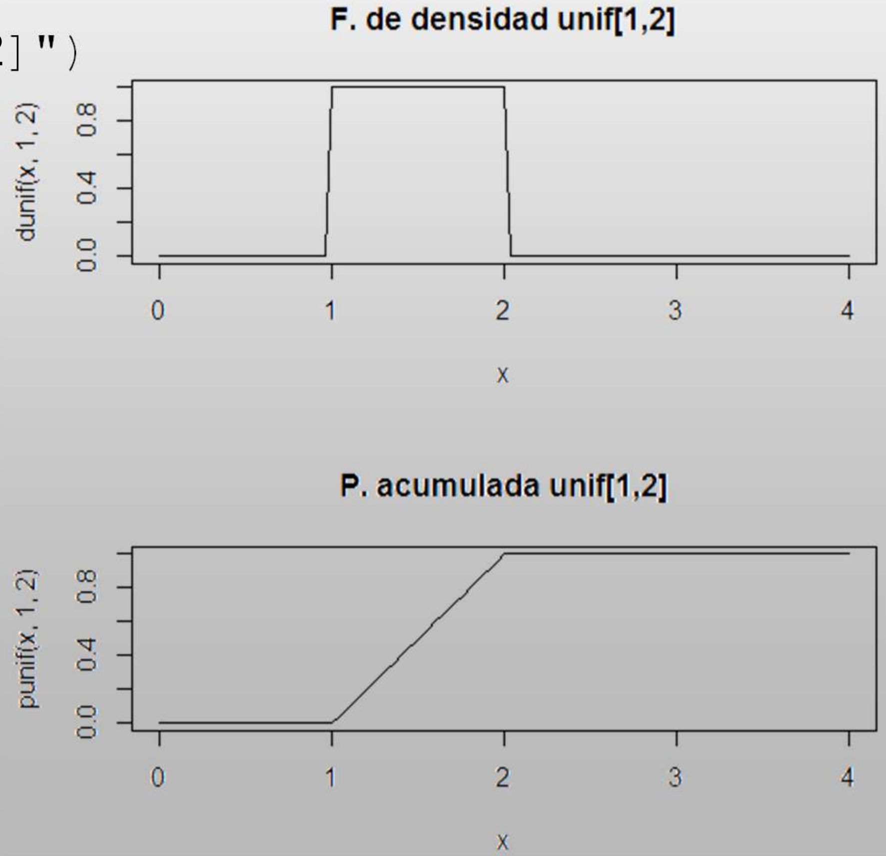
Distribución uniforme


```
>x <- runif(10^7, 1,2)
>hist(x)
>mean(x)
[1] 1.500005
>var(x)
[1] 0.08335182
```



Distribución uniforme


```
> par(mfrow=c(2,1))
> curve(dunif(x,1,2),
   from=0,to=4,
   main="F. de densidad unif[1,2]")
> curve(punif(x,1,2),
   from=0,to=4,
   main="P. acumulada unif[1,2]")
```



- Ejemplo 1:

X ~ unif[1,2], ¿P(X <= 1.5)?

$$P(X \leq 1.5) = \int_{1}^{1.5} dx = 0.5$$

```
> punif(1.5,1,2)
[1] 0.5
```

- ¿Si preguntamos al revés? Es decir, ¿cuál es el valor de z tal que $P(X \leq z)$ sea 0,5 ?

```
> qunif(0.5,1,2)
[1] 1.5
```

Distribución uniforme

- Ejemplo 2:

¿Cuál es la p que gane en la siguiente rueda de la fortuna?

$$X \sim \mathrm{unif}[0{,}360°]$$

$$P(X \leq 90°) = \frac{90 - 0}{360 - 0} = \frac{1}{4}$$

```
>punif(0,90,360)
[1]0.25
>qunif(0.25,0,360)
[190
```

Distribución exponencial

Postgrado en Modelado y Simulación

- Ya hablamos de la distribución $X \sim pois(\lambda)$ como modelo para el n de sucesos raros por unidad de tiempo
- Ahora, queremos estudiar la distribución del tiempo Y entre un suceso y el siguiente
- Así, la distribución de Y es una distribución exponencial con parámetro λ si

$$f(x) = \lambda e^{-\lambda x} \quad \text{para } x \geq 0$$

$$\mu = \frac{1}{\lambda} \qquad \sigma^2 = \frac{1}{\lambda^2}$$

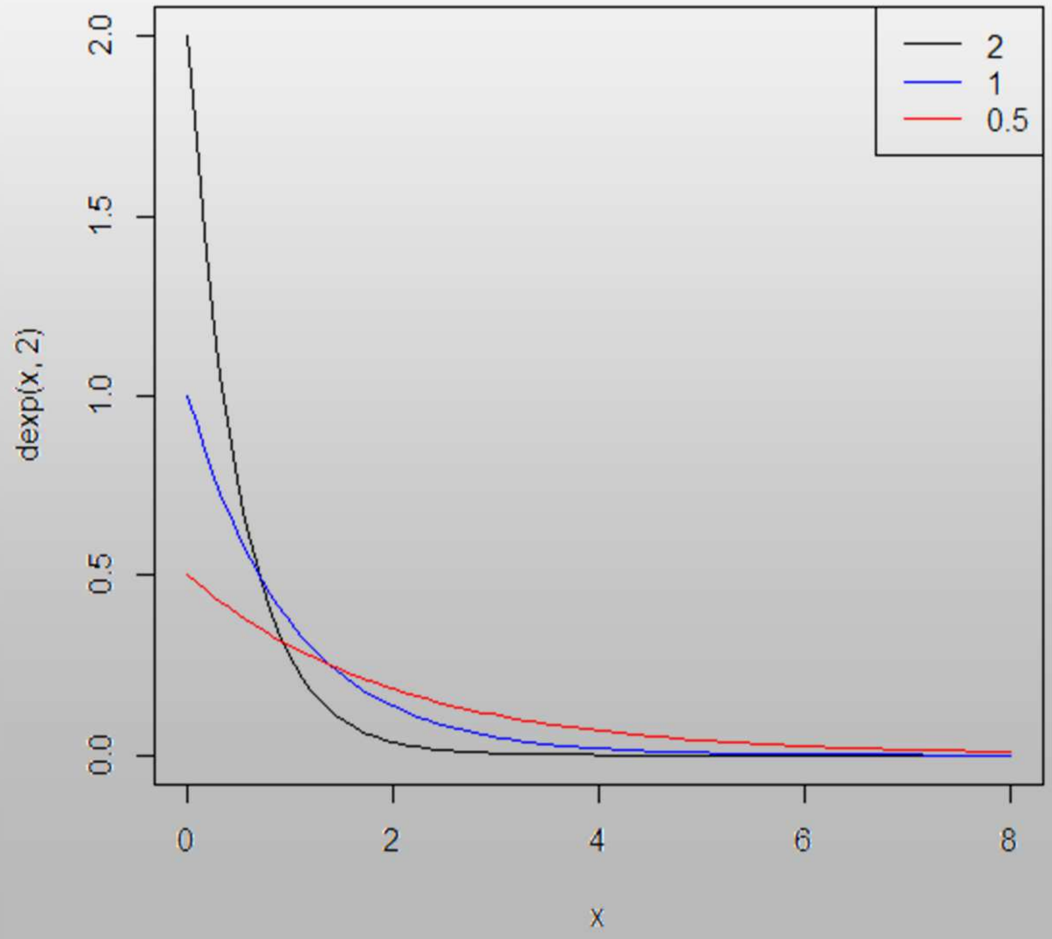
Distribución exponencial

Postgrado en Modelado y Simulación

- La distribución $X \sim exp(\lambda)$ permite modelar duración de elementos:
  - Tiempos de vida
  - Duración de huelgas
  - Periodo de desempleo
  - Tiempo entre llegadas de "clientes" a un sistema que ocurre a una tasa constante
  - Tiempo en el que falla una determinada pieza de un equipo (cuando el efecto del envejecimiento no importa y la falla se produce al azar), etc.

Distribución exponencial


```
> curve(dexp(x,2),from=0, to=8)
> curve(dexp(x,1),from=0, to=8,add=T,col="blue")
> curve(dexp(x,0.5),from=0, to=8, add=T, col="red")
> legend("topright",c("2","1","0.5"),lwd=1,col=c("black","b
  lue","red"))
```



Distribución exponencial

- Ejemplo 1: Supongamos que, en promedio, hay 10 incendios serios al año en Mérida
- Suponemos que el número de incendios por año tiene una distribución *pois(10)*

a) ¿Cuál es el tiempo promedio entre incendios?

$E[X]$ = 1/10 de año = 365/10 = 36.5 días

b) Hallar p que después del último incendio, pasen más de dos semanas hasta el siguiente

$$P(T > 14) = \int_{14}^{\infty} \frac{10}{365} e^{-\frac{10}{365}t} dt$$

$$= e^{-\frac{10}{365}*14}$$

$$\approx 0.681$$

```
>pexp(10/365,14,lower.tail=F)
[1] 0.6814301
```

Distribución exponencial

- Ejemplo 2, la magnitud de los terremotos en una región puede representarse como una exponencial λ = 0.42 en escala de Ritcher. Calcule la probabilidad de que un terremoto que azote la región:

a) Rebase los 3 grados en la escala de Richter:

$$P(X \geq 3) = \int_x^\infty \lambda e^{-\lambda x}\, dx = e^{-0.42*3} = 0.2865$$

```
>pexp(0.42,3,lower.tail=F)
[1] 0.283654
```

b) Esté entre los 2 y 3 grados en la escala de Ritcher:

$$P(2 \leq X \leq 3) = P(X \leq 3) - P(X \leq 2)$$

```
>pexp(0.42,3) - pexp(0.42,2)
[1] 0.1480565
```

Distribución exponencial

- Ejemplo 3, el tiempo entre llegadas sucesivas de carros a un punto en la carretera se distribuye como una exponencial con $\lambda = 0.01$ segundos.

a) Hallar la media y varianza de T

$$E[T] = 1/0.01 = 100 \text{ y } V[T] = 1/0{,}01^2 = 10000$$

b) Una persona empieza a cruzar la calle inmediatamente después que pasa un carro. Si tarda 50 segundos en cruzar la calle, ¿Cuál es la p que lo atropelle el siguiente carro que pasa?

```
> pexp(0.01, 50, lower.tail = F)
[1] 0.3934693
```

c) ¿Cuál es la p que no pase ningún carro en un minuto?

Sea X el número de carros que llegan en un minuto. La distribución de X es Poisson:

$$X \sim Poisson(60 \times 0{,}01) = Poisson(0{,}6)$$

$$P(X = 0) = \frac{0.6^0 e^{-0.6}}{0!} \approx 0.549$$

```
> ppois(0,0.6)
[1] 0.5488116
```

Distribución Weibull

- La distribución exponencial se caracteriza por no poseer “memoria”
- La p de observar un evento en cualquier intervalo de tiempo no depende del tiempo que ha pasado hasta ese intervalo.
- Esto es adecuado cuando los eventos aparecen al azar pero no cuando hay deterioro (i.e., desgaste de una pieza, duración de huelgas, período de desempleo, etc.) del sistema involucrado.

Distribución Weibull

Postgrado en Modelado y Simulación

- Una generalización es la distribución de Weibull
- Si $X \sim Weib(\lambda, m)$

$$f(x) = \frac{\lambda}{m}\left(\frac{x}{m}\right)^{\lambda-1} e^{\left(\frac{-x}{m}\right)^{\lambda}} \quad para\ x \geq 0$$

$$\mu = m\Gamma(1+\frac{1}{\lambda})\ \sigma^2 = m^2\left[\Gamma\left(1+\frac{2}{\lambda}\right) - \Gamma\left(1+\frac{1}{\lambda}\right)^2\right]$$

$$\Gamma(p) = \int_0^{\infty} x^{p-1}e^{-x}\,dx \qquad \text{para} \quad p > 0$$

$$\Gamma(p) = (p-1)\Gamma(p-1); \Gamma(1) = 1; \Gamma\left(\frac{1}{2}\right) = \sqrt{(\pi)}$$

$$Si\ n\ es\ entero\ \Gamma(n) = (n-1)$$

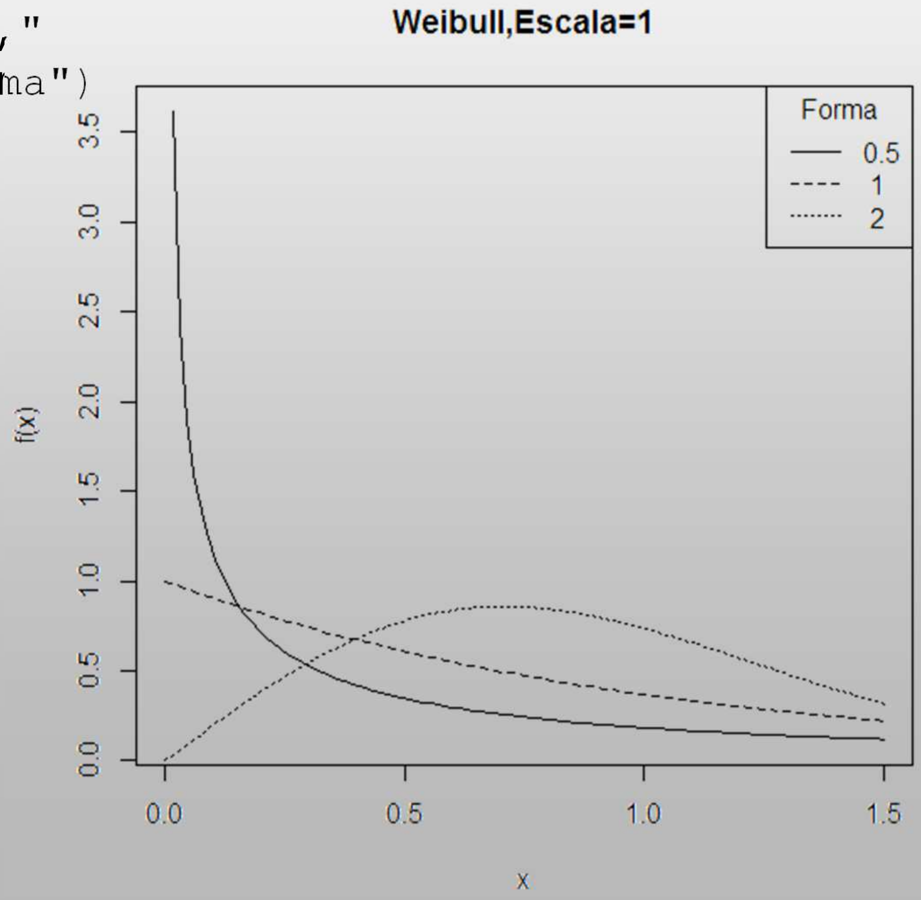
Distribución Weibull


```
> curve(dweibull(x,shape=0.5,scale=1),from=0,
     to= 1.5,ylab="f(x)",main="weibull,escala=1")
> curve(dweibull(x,shape=1,scale=1),from=0,
     to= 1.5,lty=2,add=t)
> curve(dweibull(x,shape=2,scale=1),from=0,
     to= 1.5,lty=3,add=t)
> legend("topright",c("0.5"," 1","
     2"),lty=c(1,2,3),title="forma")
```

En R:

shape = λ

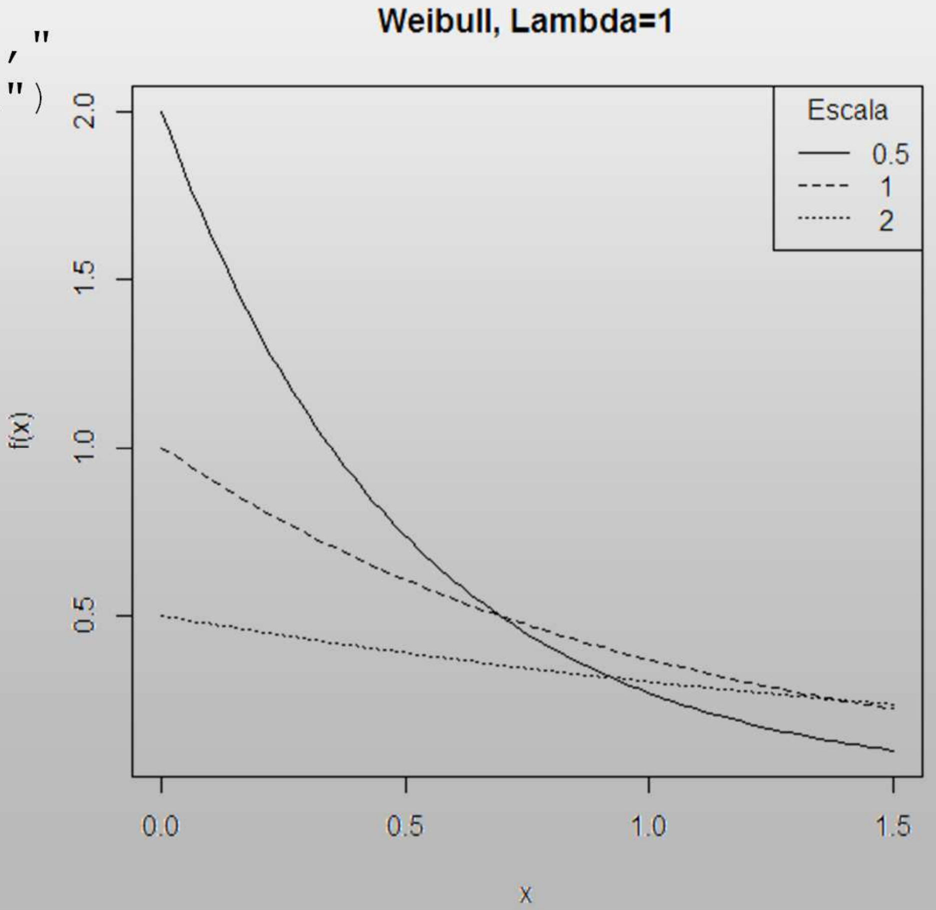
scale = m



Distribución Weibull

```
> curve(dweibull(x,shape=1,scale=0.5),from=0,
     to= 1.5,ylab="f(x)",main="Weibull, Lambda=1")
> curve(dweibull(x,shape=1,scale=1),from=0,
     to= 1.5,lty=2,add=T)
> curve(dweibull(x,shape=1,scale=2),from=0,
     to= 1.5,lty=3,add=T)
> legend("topright",c("0.5"," 1","
  2"),lty=c(1,2,3),title="Escala")
```

En R:
shape = λ
scale = m



Distribución Pareto

- En los trabajos pioneros para describir la distribución de la renta, Vilfredo Pareto, tras estudiar la renta en numerosos países, observó que al menos para valores grandes:

$$\ln(N) \approx \ln(A) - \alpha x$$

donde A y $\alpha > 0$ son parámetros y N es el número de individuos que perciben una renta mayor o igual que x

Distribución Pareto

Postgrado en Modelado y Simulación

- Sea X una variable aleatoria, la función de Pareto(α, x_0) es

$$f(x) = \begin{cases} 1 - \left(\frac{x_0}{x}\right)^{\alpha} & si \quad x \geq x_0 \\ 0 & si \quad x < x_0 \end{cases}$$

$$E[X] = \frac{\alpha x_0}{\alpha - 1}$$

$$Var[X] = \frac{\alpha {x_0}^2}{(\alpha - 2)(\alpha - 1)^2}$$

Distribución Gamma

- Si $X \sim gamma$

$$f(x) = \frac{1}{\Gamma(m)} \lambda^m x^{m-1} e^{-\lambda x} para x \geq 0; \lambda y m \geq 0$$

$$\mu = \frac{m}{\lambda} \sigma^2 = \frac{m}{\lambda^2}$$

- En R

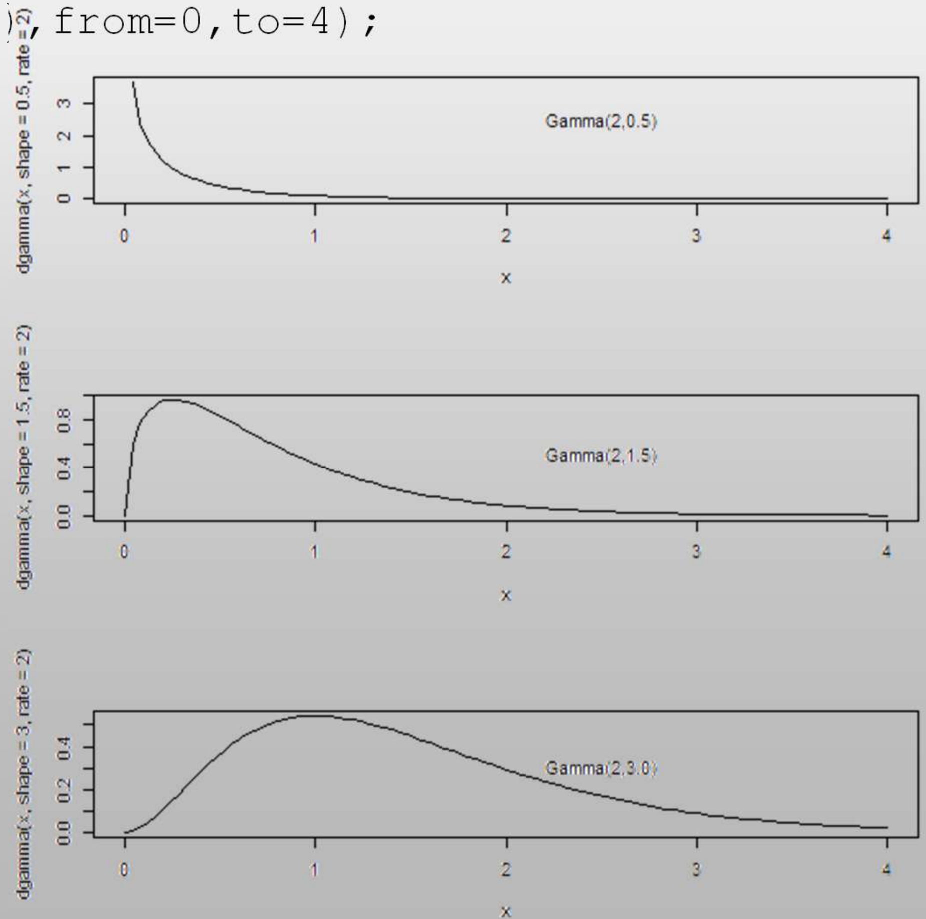
  `>shape` (forma) = m

  `>rate` (tasa) = λ

- Si m es entero, se le llama distribución Earlang

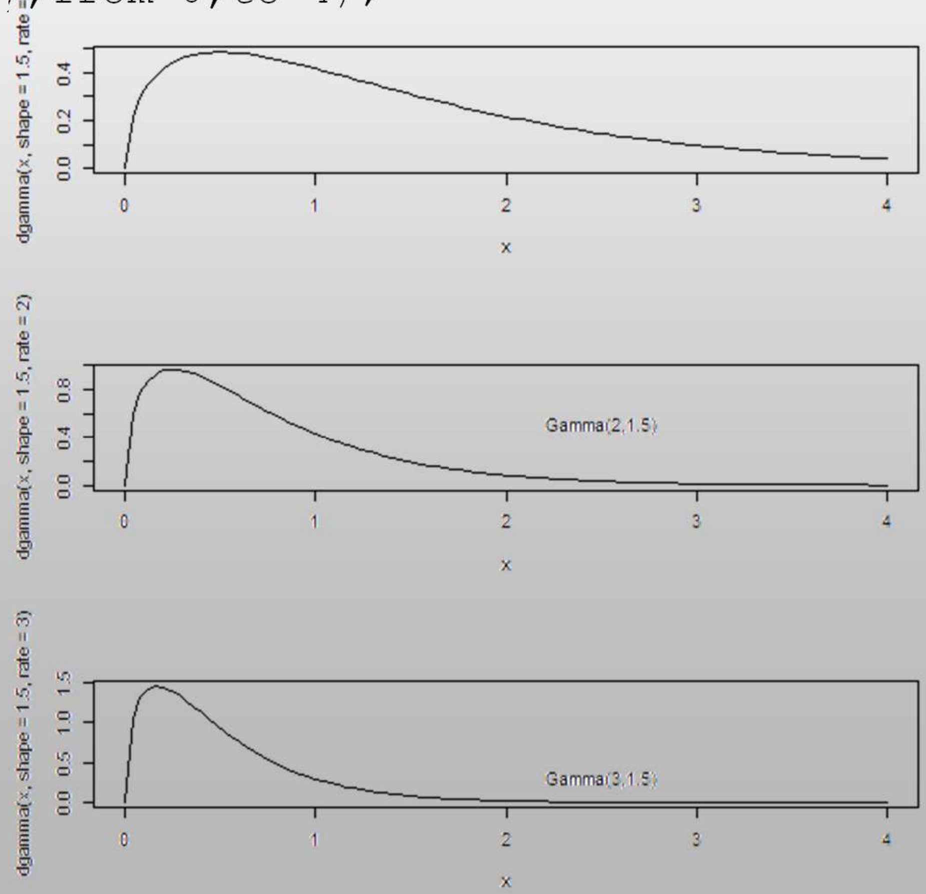
Distribución Gamma


```
> par(mfrow=c(3,1))
> curve(dgamma(x,shape=0.5,rate=2),from=0,to=4);
  text(2.5,2.5,"Gamma(2,0.5)")
> curve(dgamma(x,shape=1.5,rate=2),from=0,to=4);
  text(2.5,0.5,"Gamma(2,1.5)")
> curve(dgamma(x,shape=3.0,rate=2),from=0,to=4);
  text(2.5,0.3,"Gamma(2,3.0)")
```



Distribución Gamma


```
> par(mfrow=c(3,1))
> curve(dgamma(x,shape=1.5,rate=1),from=0,to=4);
  text(2.5,2.5,"Gamma(1,1.5)")
> curve(dgamma(x,shape=1.5,rate=2),from=0,to=4);
  text(2.5,0.5,"Gamma(2,1.5)")
> curve(dgamma(x,shape=1.5,rate=3),from=0,to=4);
  text(2.5,0.3,"Gamma(3,1.5)")
```



Distribución Gamma

- Las precipitaciones durante un mes de sequía en una cierta región tienen una distribución gamma con parámetros m=1.6 y λ=0.5. Calcule la media y la varianza de esta variable

$$\mu = \frac{m}{\lambda} = 1.6/0.5 = 3.2$$

$$\sigma^2 = \frac{m}{\lambda^2} = 1.6/(0.5^2) = 6.4$$

- ¿Cuál es la probabilidad de que llueva menos que la media?

Distribución normal

- Cuando se toma el promedio de un número "suficientemente" grande de una muestra aleatoria iid la distribución del promedio luce como una distribución normal
- También se usa para modelizar errores de medición
- Si $X \sim \text{norm}(\mu,\sigma^2)$ entonces:

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

- $Z \sim \text{norm}(0,1)$ se le conoce como normal estándar. F(z) se denota con la letra φ y a f(z) como Φ
- La función de probabilidad acumulada no puede obtenerse analíticamente de allí las famosas tablas de la normal

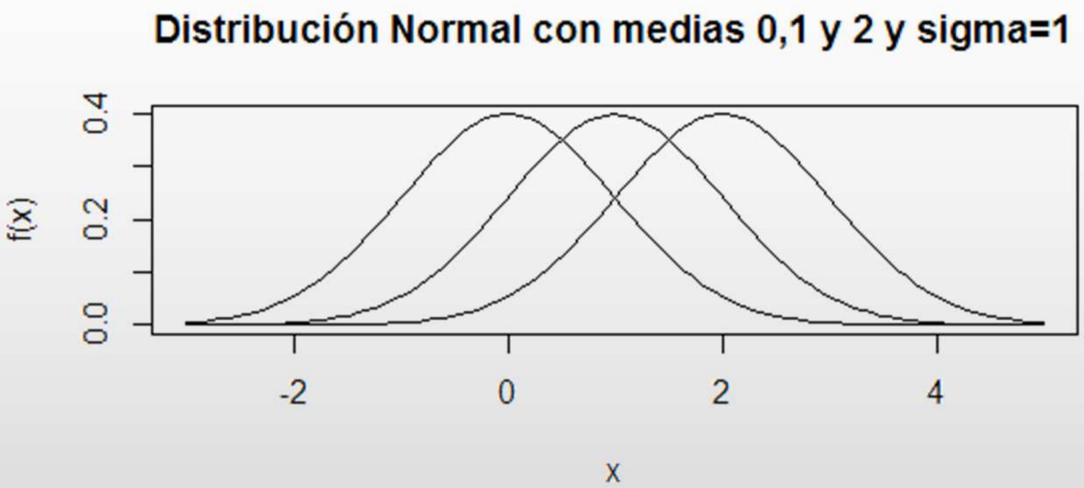
Distribución normal

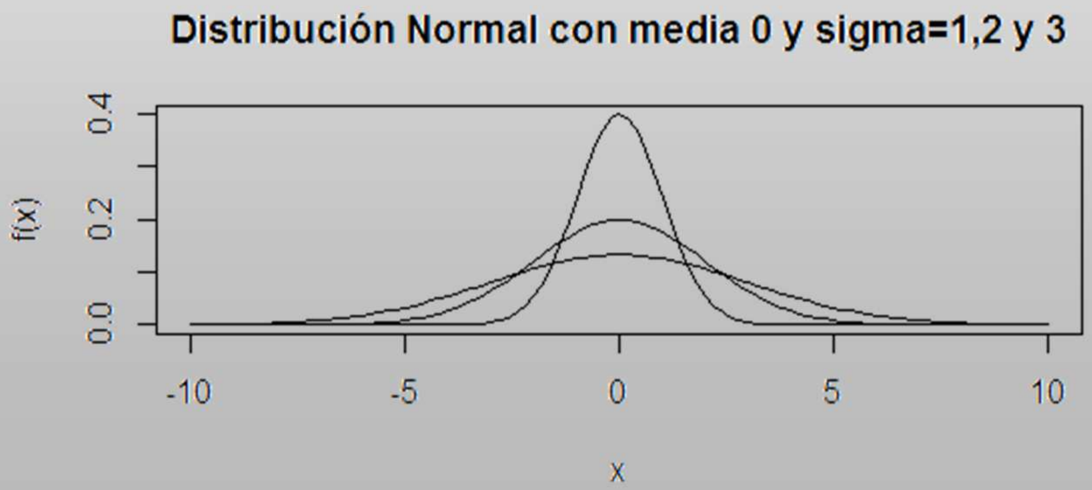
Postgrado en Modelado y Simulación

```
> par(mfrow=c(2,1))
> curve(dnorm(x,mean=0,sd=1),from=-3,
  to=5,ylab="f(x)",main="Distribución    Normal    con
  medias 0,1 y 2 y sigma=1")
> curve(dnorm(x,mean=1,sd=1),from=-3,     to=5,add=T);
  curve(dnorm(x,mean=2,sd=1),from=-3, to=5,add=T)
> curve(dnorm(x,mean=0,sd=1),from=-10,
  to=10,ylab="f(x)",main="Distribución    Normal    con
  media 0 y sigma=1,2 y 3")
> curve(dnorm(x,mean=0,sd=2),from=-10,   to=10,add=T);
  curve(dnorm(x,mean=0,sd=3),from=-10, to=10,add=T)
```

Distribución normal

Postgrado en Modelado y Simulación





Distribución normal

Postgrado en Modelado y Simulación

- Los resultados de un examen de admisión (escala 1-100) poseen una distribución normal con media 75 y desviación 10

a) ¿Qué fracción de los resultados está entre 80 y 90?

$P(80 < x < 90) = p(x < 90) - p(x < 80)$

```
> pnorm(90,75,10) – pnorm(80,75,10)
[1] 0.2417
```

b) ¿Cuál es la probabilidad de obtener una nota superior a 90?

```
> pnorm(90,75,10,lower.tail=F)
> 1 – pnorm(90,75,10)
> pnorm(1.5,0,1)       #estandarizando
[1]0.0668072
```

c) Como el cupo es reducido, ¿Cuál es la nota de corte si quiero que solo ingresen aquellos con notas superiores al 90% de todos los resultados?

```
> qnorm(0.90,75,10)
[1] 87.81552
```

OTRAS DISTRIBUCIONES

- La distribución X^2: suma de normales estándares al cuadrado. Sirve para pruebas de bondad del ajuste y varianza
- La distribución *t* de student: similar a la normal pero con colas más anchas. Se usa para calcular IC cuando la varianza es desconocida
- La distribución lognormal: modeliza tiempos de tareas, cantidades que son el producto de un número grande de variables