UNIVERSITA' DEGLI STUDI DI VERONA

LABORATORIO DI PROBABILITA' E STATISTICA

Docente: Bruno Gobbi

5 – VARIABILI CASUALI DISCRETE

LA VARIABILE BINOMIALE

Sia $n \in \mathbf{N}$ e sia $k \leq n$. La probabilità di osservare k successi in n prove, ciascuna con probabilità $p \in (0;1)$ è

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

LA VARIABILE BINOMIALE

I momenti della variabile binomiale sono:

- Media: $\mu = np$
- Varianza: $\sigma^2 = npq$
- Scarto quadratico medio: $\sigma = \sqrt{npq}$

LA VARIABILE BINOMIALE

In R si definiscono quattro funzioni per la variabile binomiale:

- **dbinom()** calcola la densità di probabilità
- **pbinom()** è la funzione di probabilità cumulata
- **qbinom()** è l'inversa della probabilità cumulata
- **rbinom()** per creare dei valori random generati da una variabile casuale binomiale

ESEMPIO DI VARIABILE BINOMIALE

Ipotizziamo che la probabilità di passare l’esame di statistica sia del 70%. Supponendo che 5 studenti si presentino all’appello, descrivere con una opportuna variabile casuale le probabilità che gli studenti vengano promossi.

ESEMPIO DI VARIABILE BINOMIALE

p = 0,7

n = 5

k = 0, 1, 2, 3, 4, 5

$$\binom{n}{k} p^k (1-p)^{n-k}$$

ESEMPIO DI VARIABILE BINOMIALE

k	P(k)
0	0,24%
1	2,84%
2	13,23%
3	30,87%
4	36,02%
5	16,81%
TOT	**100%**

LA FUNZIONE dbinom(k, n, p)

 # CREO IL VETTORE DEI k

> k=c(0:5)

> k

[1] 0 1 2 3 4 5

 # CALCOLO LE PROBABILITA' DELLA BINOMIALE CON LA FUNZIONE dbinom

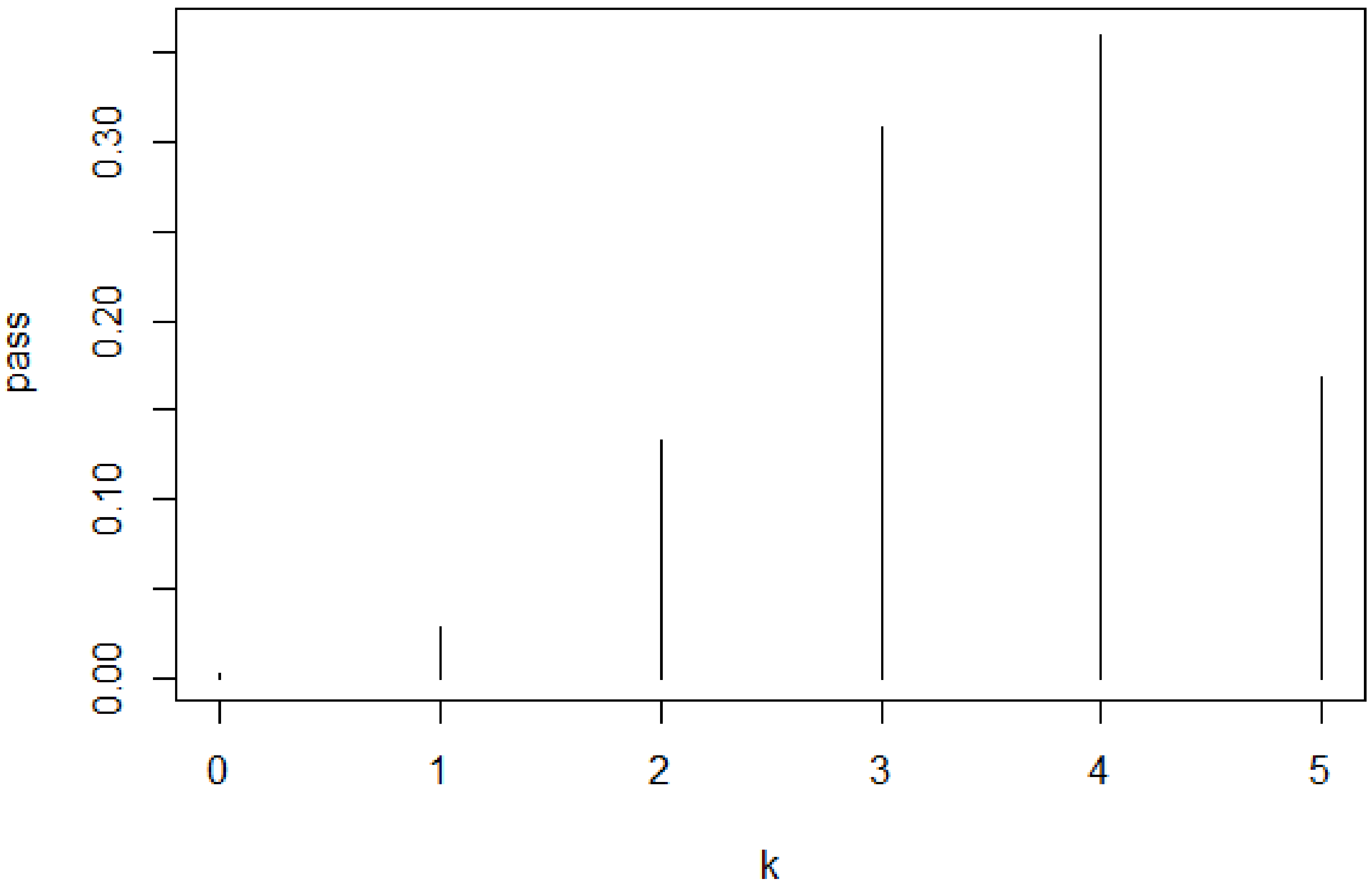
> pass=dbinom(k, 5, 0.7) **# dbinom(k, n, p)**

> pass

[1] 0.00243 0.02835 0.13230 0.30870 0.36015 0.16807

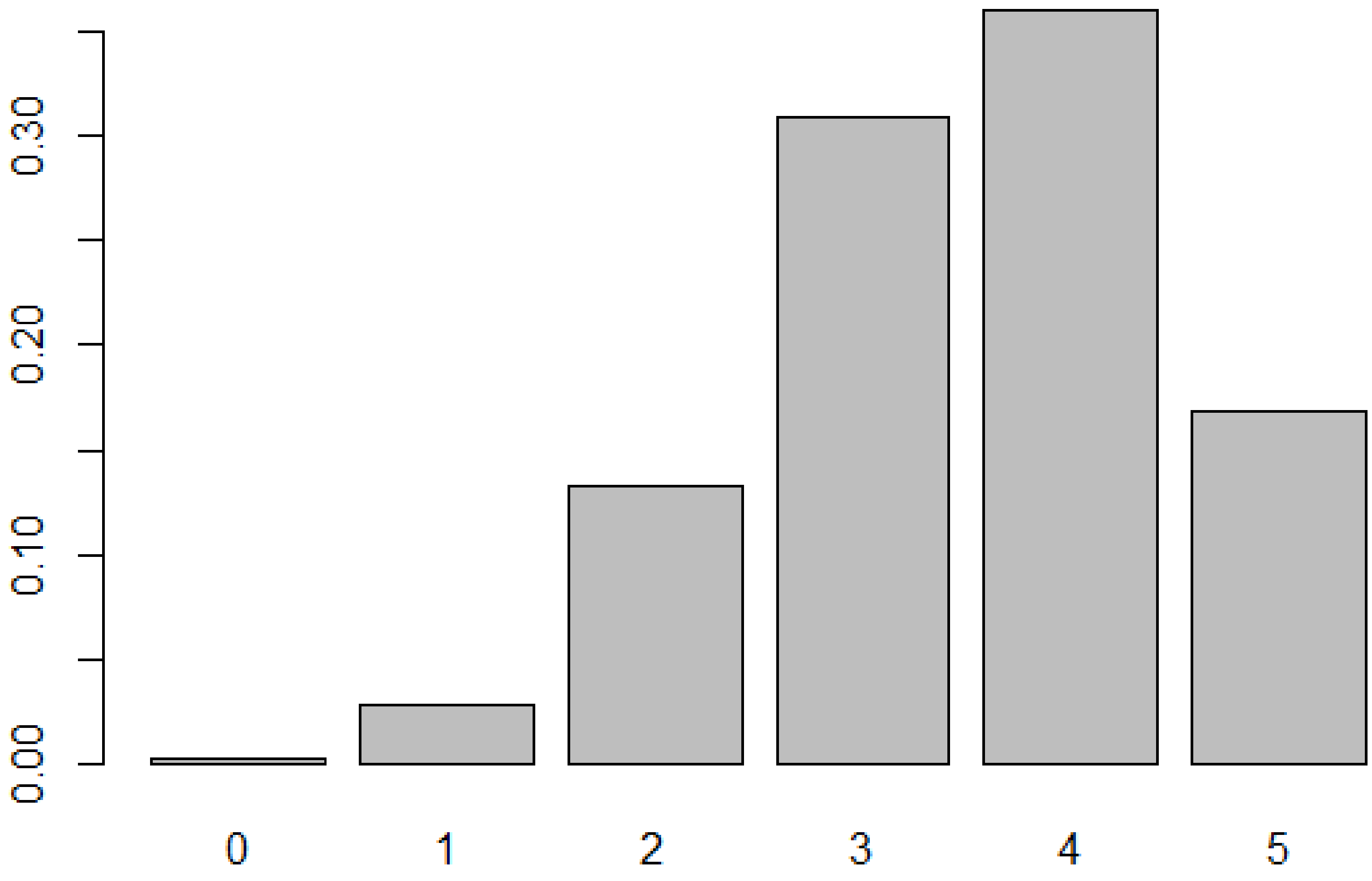
DISEGNO IL GRAFICO DELLA VARIABILE

> plot(k, pass, 'h') **# ‘h’ CREA LE LINEE**



PER UN GRAFICO PIU' ELEGANTE

```
> barplot(pass, names.arg=k)
```



ESEMPIO FIGLI MASCHI

La probabilità di avere figli maschi è di 0.52 (alla nascita ci sono leggermente più maschi che femmine, si equivalgono intorno ai 40 anni e in seguito sono di più le femmine)

Descrivere con una opportuna variabile casuale quali sono le probabilità di avere dei maschi in una famiglia con 4 figli.

ESEMPIO FIGLI MASCHI

 # CREO IL VETTORE DEI k

```
> k=c(0:4)
> k
[1] 0 1 2 3 4
```

 # CALCOLO LE PROBABILITA' CON dbinom

```
> figlimaschi=dbinom(k, 4, 0.52)
> figlimaschi
[1] 0.05308416 0.23003136 0.37380096
0.26996736 0.07311616
```

DISEGNO IL GRAFICO DELLA VARIABILE

```
> plot(k, figlimaschi, 'h')
```

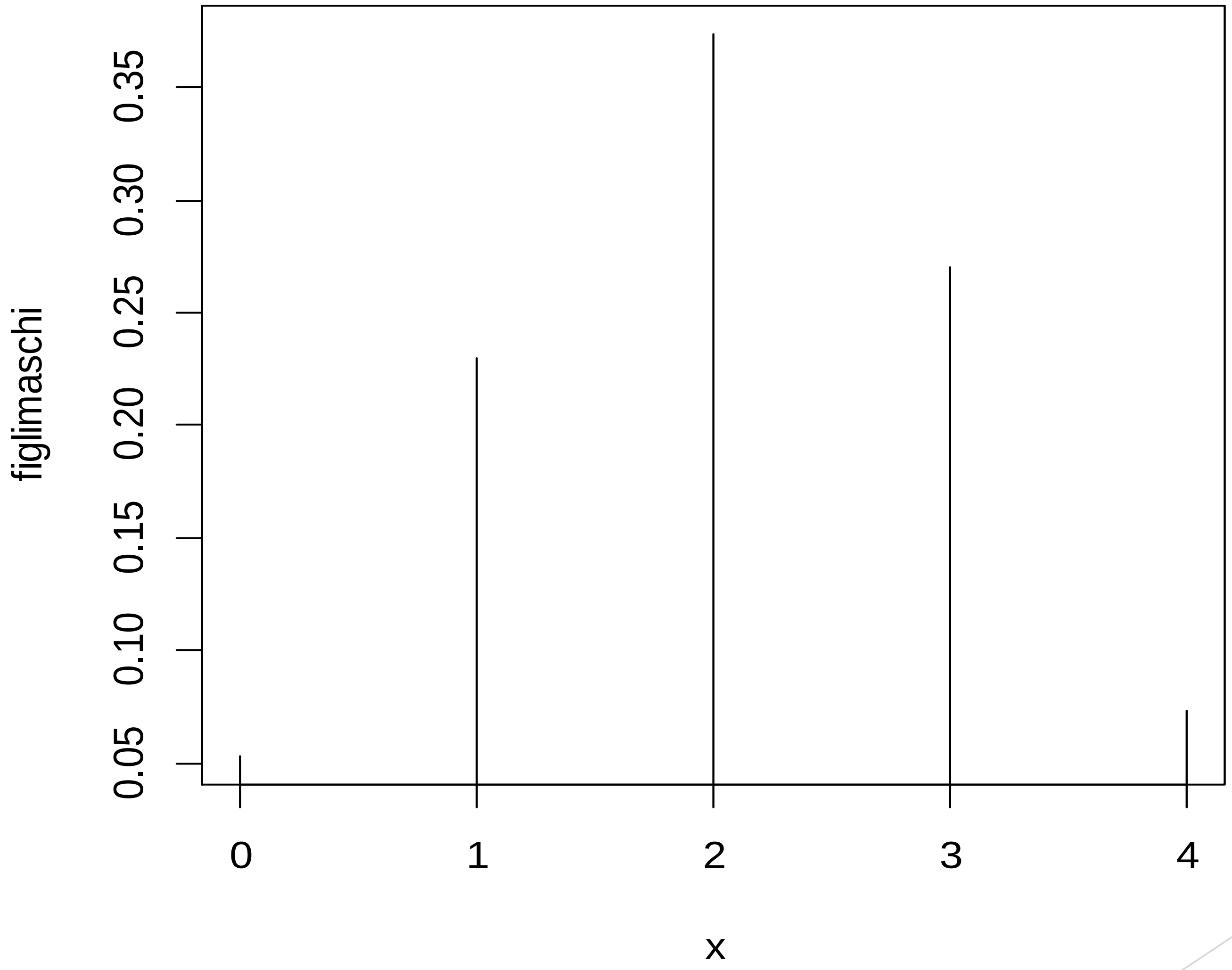
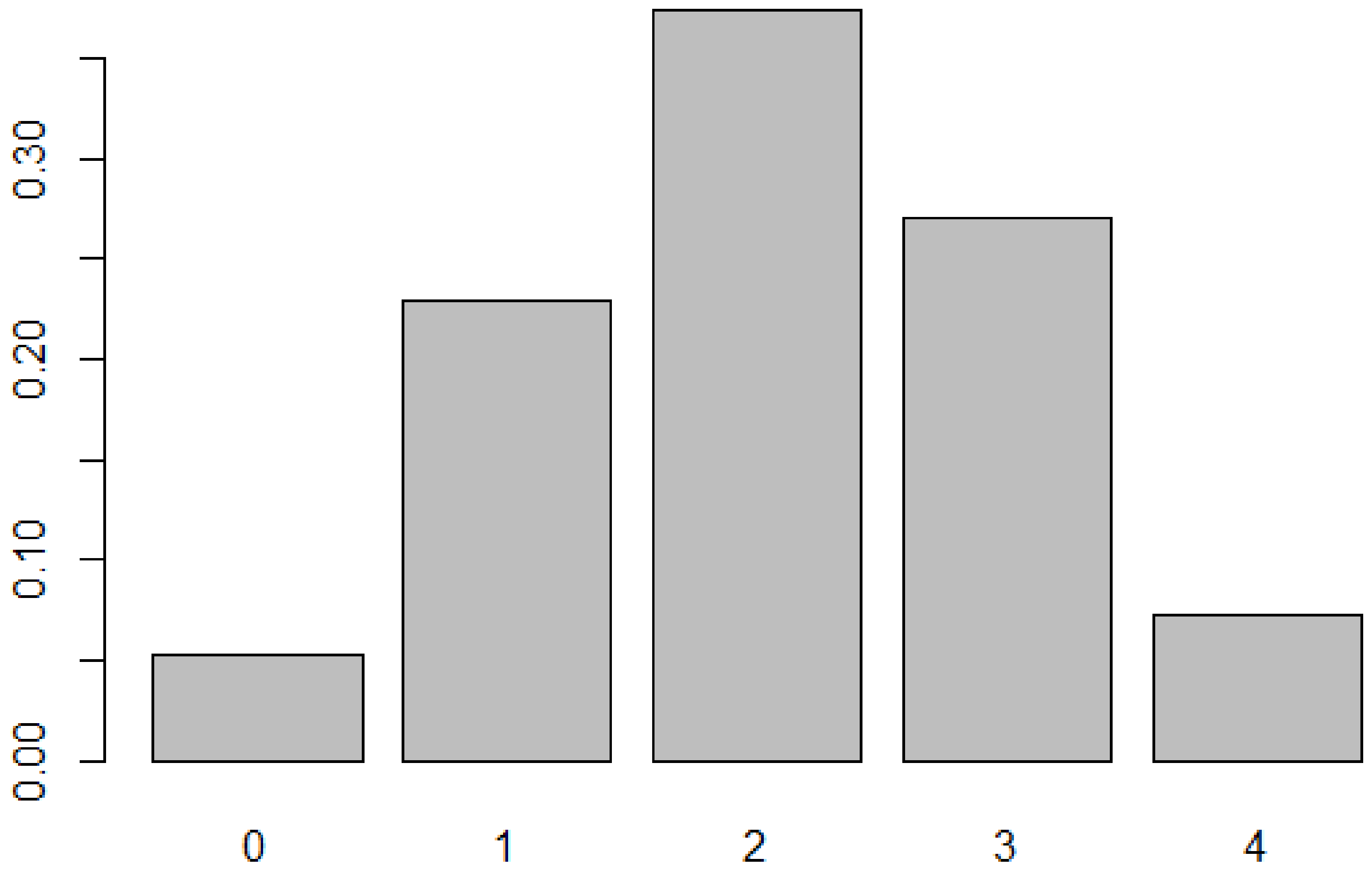


GRAFICO FIGLI MASCHI

```
> barplot(figlimaschi, names.arg=k)
```



ESEMPIO FIGLIE FEMMINE

E se volessimo calcolare le corrispondenti probabilità di avere una figlia, sempre in una famiglia con 4 bambini?

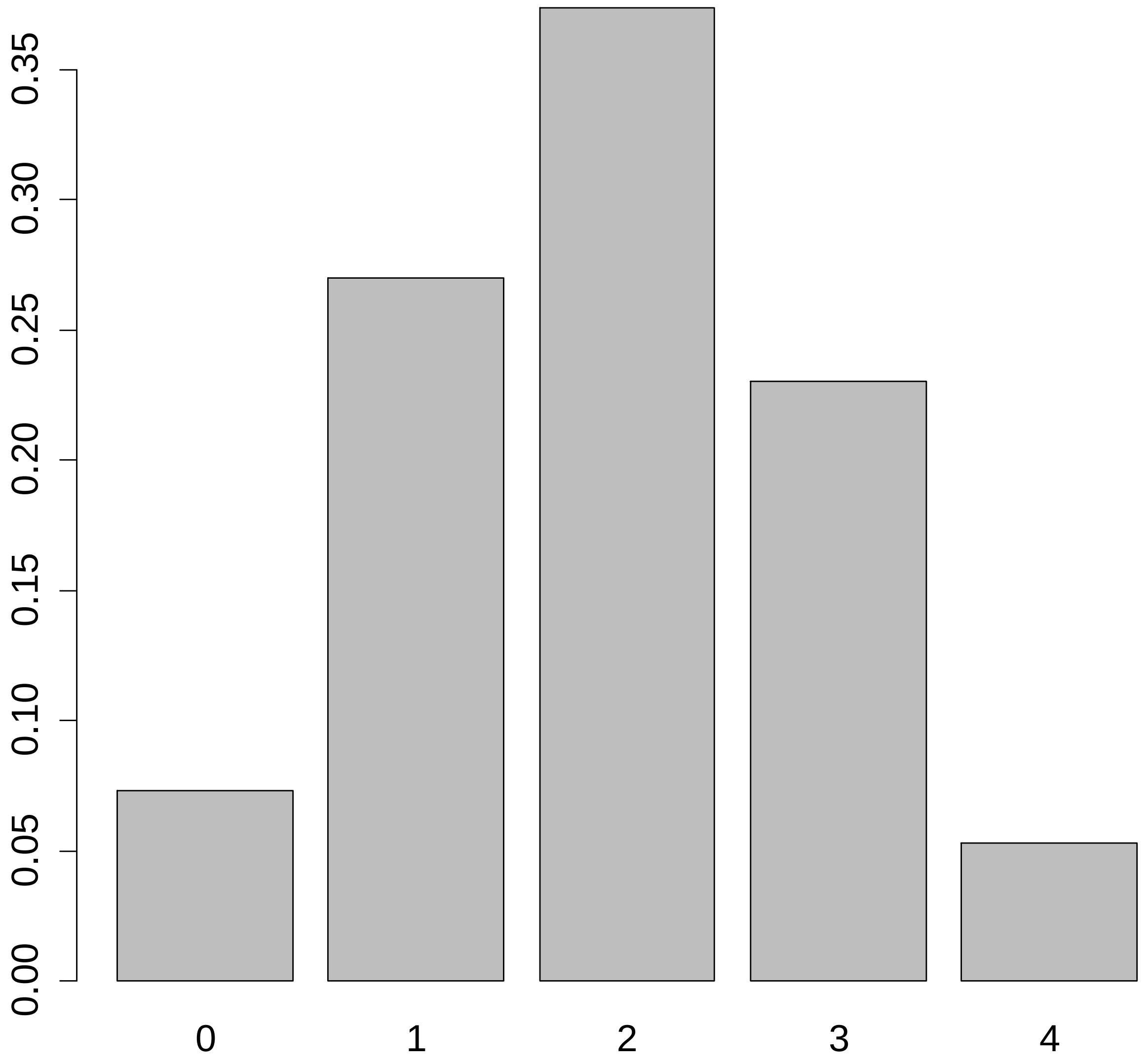
ESEMPIO FIGLIE FEMMINE

 # CALCOLO LE PROBABILITA' CON dbinom

```
> figliefemmine=dbinom(k, 4, 0.48)
> figliefemmine
[1] 0.07311616 0.26996736 0.37380096
0.23003136 0.05308416
```

GRAFICO FIGLIE FEMMINE

> barplot(figliefemmine, names.arg=k)



ESEMPIO LANCIO DI UNA MONETA

Supponendo di lanciare 10 volte in aria una moneta (non truccata), descrivere la probabilità di ottenere testa con una opportuna variabile casuale.

ESEMPIO LANCIO MONETA

 # CREO IL VETTORE DEI k

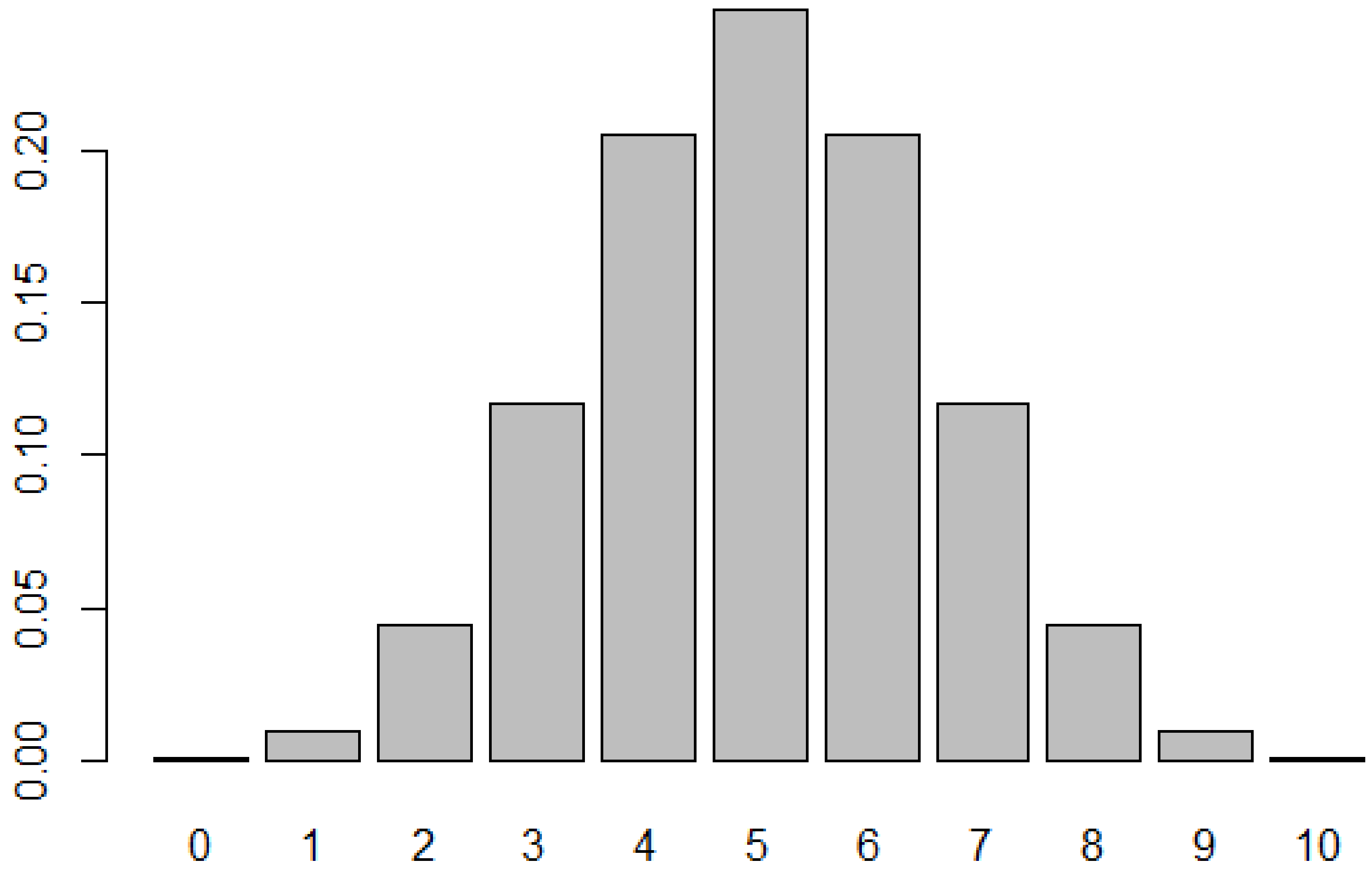
```
> k=c(0:10)
```

 # CALCOLO LE PROBABILITA' CON dbinom

```
> moneta=dbinom(k, 10, 0.5)
> moneta
 [1] 0.0009765625 0.0097656250 0.0439453125 0.1171875000 0.2050781250 0.2460937500
 [7] 0.2050781250 0.1171875000 0.0439453125 0.0097656250 0.0009765625
```

GRAFICO MONETA

> barplot(moneta, names.arg=k)



ESEMPIO MONETA TRUCCATA

E se invece la moneta fosse truccata, tale per cui la probabilità che esca testa sia leggermente superiore a quella che esca croce, ad es. 0.54?

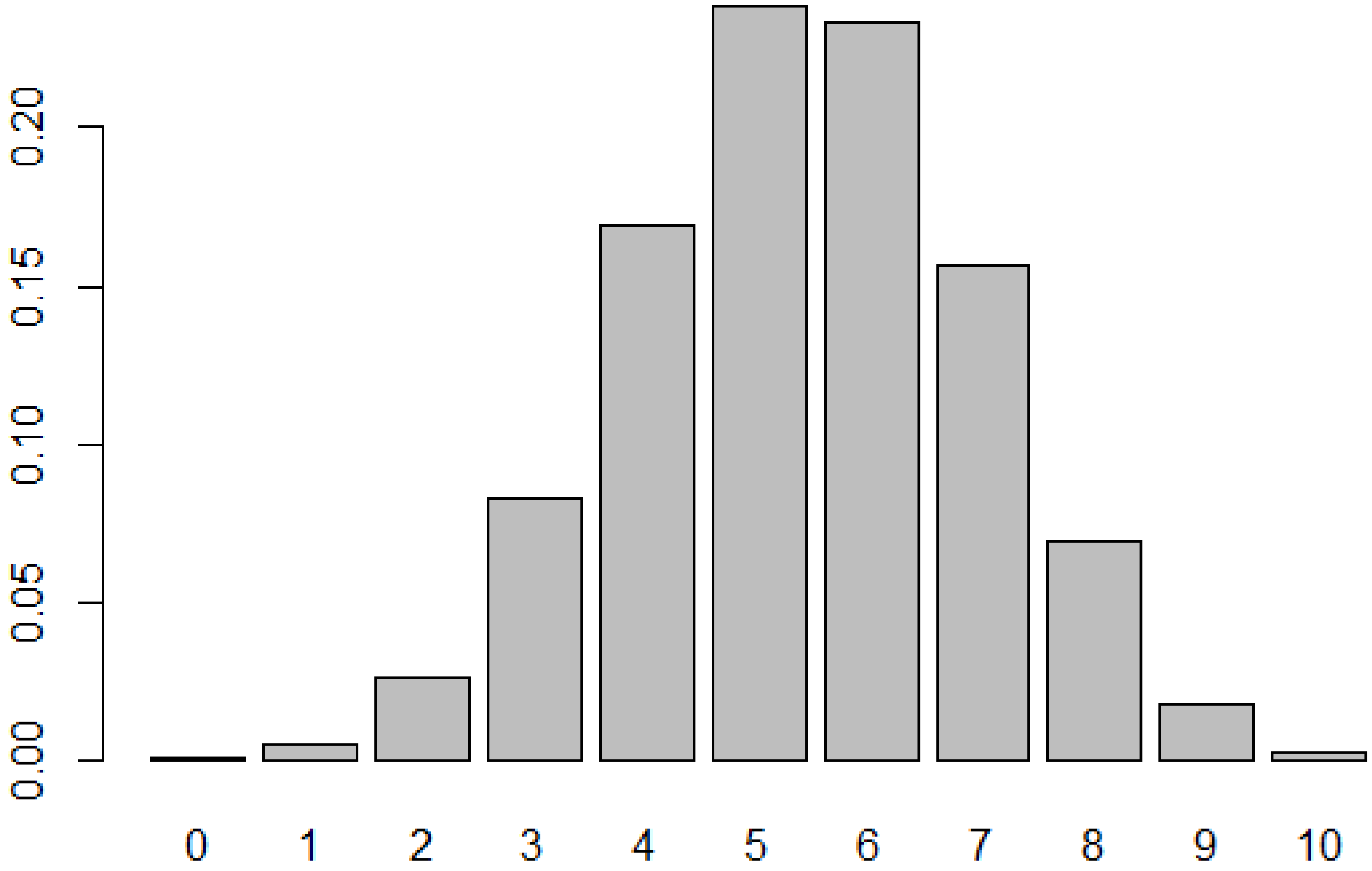
ESEMPIO MONETA TRUCCATA

 # CALCOLO LE PROBABILITA' CON dbinom

```
> monetatruc=dbinom(k, 10, 0.54)
> monetatruc
 [1] 0.0004242075 0.0049798269 0.0263064768 0.0823507099 0.1691770018 0.2383189069
 [7] 0.2331380611 0.1563907491 0.0688459276 0.0179598072 0.0021083252
```

GRAFICO MONETA TRUCCATA

```
> barplot(monetatruc, names.arg=k)
```



Asimmetria
a destra

ESEMPIO QUIZ CON RISPOSTA A CASO

Supponiamo che Tizio debba fare un test con 30 domande, ciascuna domanda con 3 risposte; immaginiamo che Tizio risponda a caso.

Descrivere con una opportuna variabile casuale, indicandone anche media e varianza, e calcolare:

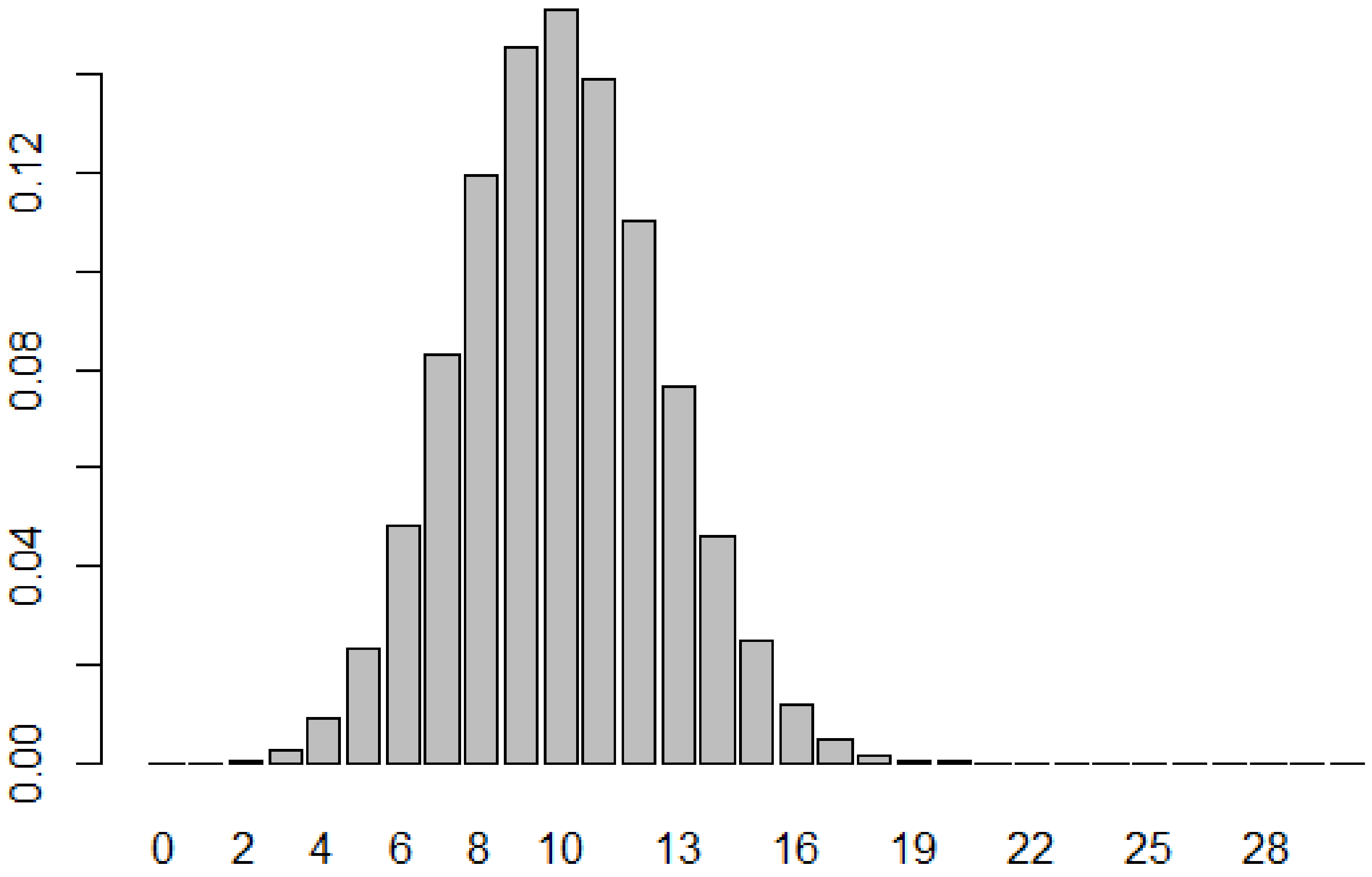
- La probabilità di rispondere correttamente a 18 domande
- La probabilità di rispondere correttamente a una sola domanda
- La probabilità di rispondere correttamente a un numero di domande compreso fra 0 e 10

ESEMPIO QUIZ CON RISPOSTA A CASO

```
> k=c(0:30)
> test=dbinom(k, 30, 1/3)
> test
 [1] 5.215095e-06 7.822643e-05 5.671416e-04 2.646661e-03
8.932480e-03 2.322445e-02
 [7] 4.838427e-02 8.294446e-02 1.192327e-01 1.457288e-01
1.530152e-01 1.391048e-01
[13] 1.101246e-01 7.624011e-02 4.628864e-02 2.468727e-02
1.157216e-02 4.765007e-03
[19] 1.720697e-03 5.433780e-04 1.494289e-04 3.557832e-05
7.277384e-06 1.265632e-06
[25] 1.845713e-07 2.214856e-08 2.129669e-09 1.577533e-10
8.451068e-12 2.914161e-13
[31] 4.856936e-15
```

GRAFICO QUIZ CON RISPOSTA A CASO

```
> barplot(test, names.arg=k)
```



 # Media: $\mu = np = 30 * 1/3 = 10$
 # Varianza: $\sigma^2 = npq = 30 * 1/3 * 2/3 = 6,6666$

ESEMPIO QUIZ CON RISPOSTA A CASO

 # CALCOLO LA PROBABILITÀ DI RISPONDERE CORRETTAMENTE A 18 DOMANDE

```
> test18=dbinom(18, 30, 1/3)
> test18
[1] 0.001720697
```

 # LA PROBABILITA' CHE TIZIO RISPONDA CORRETTAMENTE A 18 DOMANDE RISPONDENDO A CASO E' DELLO 0,172%

ESEMPIO QUIZ CON RISPOSTA A CASO

 # CALCOLO LA PROBABILITÀ DI RISPONDERE CORRETTAMENTE A 1 SOLA DOMANDA

```
> test1=dbinom(1, 30, 1/3)
> test1
[1] 7.822643e-05
```

 # LA PROBABILITA' CHE TIZIO INDOVINI UNA SOLA DOMANDA E' DELLO 0,0007822%

LA FUNZIONE pbinom

 # CALCOLO LA PROBABILITÀ DI RISPONDERE CORRETTAMENTE A UN NUMERO DI DOMANDE COMPRESO FRA 0 E 10

```
> test10p=pbinom(10, 30, 1/3)
> test10p
[1] 0.5847596
```

 # IN QUESTO CASO SI USA LA FUNZIONE pbinom PER LA CUMULATA DELLE PROBABILITA'

ESEMPIO QUIZ CON RISPOSTA A CASO

Sui dati dell’esercizio precedente, calcolare:

- La probabilità di rispondere correttamente a più di 10 domande (da 11 a 30)
- La probabilità di rispondere correttamente a un numero di domande compreso fra 7 e 12
- Il valore mediano

ESEMPIO QUIZ CON RISPOSTA A CASO

 # CALCOLO LA PROBABILITÀ DI RISPONDERE CORRETTAMENTE A PIU' DI 10 DOMANDE (DA 11 A 30)

```
> 1-pbinom(10, 30, 1/3)
[1] 0.4152404
```

 # LA PROBABILITA' DI RISPONDERE CORRETTAMENTE A PIU' DI 10 DOMANDE E' DEL 41,524%

ESEMPIO QUIZ CON RISPOSTA A CASO

```
# OPPURE …

> 1-pbinom(10, 30, 1/3)
[1] 0.4152404
> pbinom(10, 30, 1/3, lower.tail=FALSE)
[1] 0.4152404
```

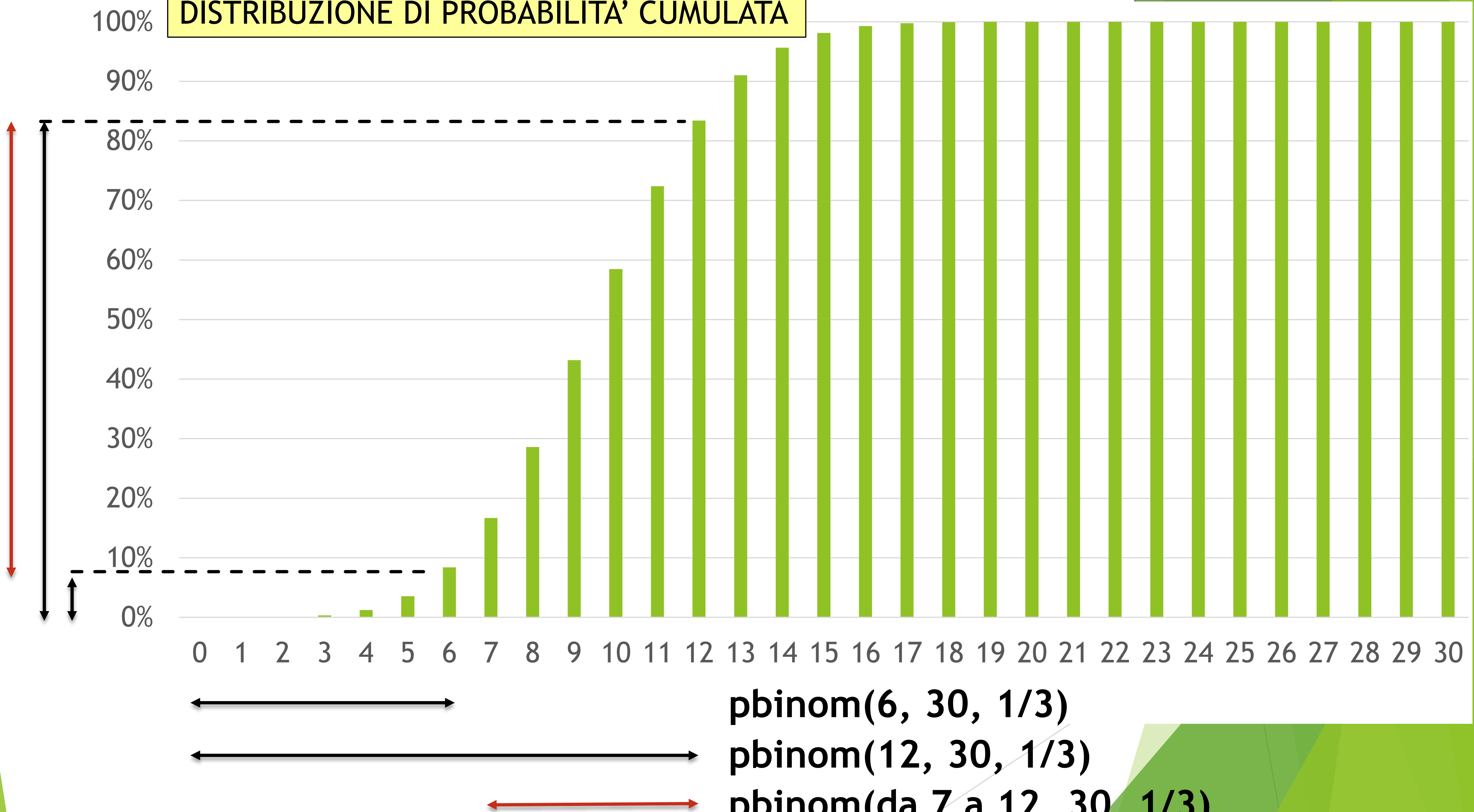
ESEMPIO QUIZ CON RISPOSTA A CASO

 # CALCOLO LA PROBABILITÀ DI RISPONDERE CORRETTAMENTE A UN NUMERO DI DOMANDE COMPRESO FRA 7 E 12

```
> test12p=pbinom(12, 30, 1/3)
> test6p=pbinom(6, 30, 1/3)
> test_da_7_a_12=test12p-test6p
> test_da_7_a_12
[1] 0.7501505
```

 # LA PROBABILITA' DI RISPONDERE CORRETTAMENTE A UN NUMERO DI DOMANDE COMPRESO FRA 7 E 12 E' DEL 75,015%

DISTRIBUZIONE DI PROBABILITA’ CUMULATA



pbinom(6, 30, 1/3)

pbinom(12, 30, 1/3)

pbinom(da 7 a 12, 30, 1/3)

LA FUNZIONE qbinom

 # CALCOLO IL VALORE MEDIANO

```
> test_mediana=qbinom(0.5, 30, 1/3)
> test_mediana
[1] 10
```

 # LA FUNZIONE qbinom(percentuale, n, p) E' L'INVERSA DELLA qbinom, RESTITUISCE IL VALORE DI *k* CORRISPONDENTE AD UNA CERTA PROBABILITA'. AD ES. A 0,5 CORRISPONDE IL VALORE CENTRALE, OSSIA LA MEDIANA.

LA FUNZIONE rbinom

 # PER OTTENERE DEI VALORI GENERATI A RANDOM CHE SEGUONO LO SCHEMA BINOMIALE SI USA:

> rbinom(*n. tentativi*, *n*, *p*)

IL "*n. tentativi*" INDICA QUANTE VOLTE SI VUOLE RIPETERE L'ESPERIMENTO OVVERO QUANTI RISULTATI VERRANNO GENERATI CASUALMENTE DA R SEGUENDO LO SCHEMA BINOMIALE CON PARAMETRI *n* E *p*.

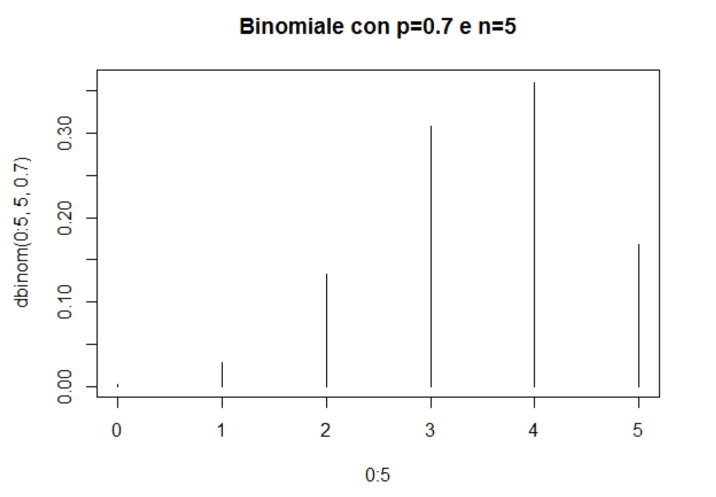
GENERAZIONE RANDOM DI BINOMIALE

 # ES. IPOTIZZANDO CHE TIZIO PROVI 5 VOLTE L'ESAME RISPONDENDO A CASO, QUALI VOTI PRENDERA'?

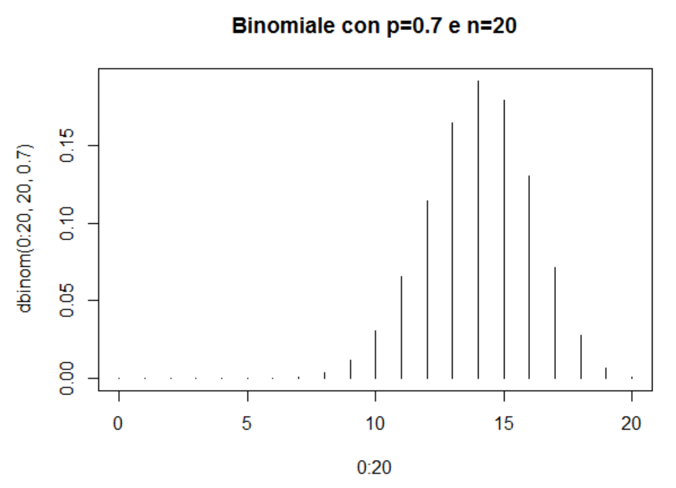
```
> rbinom(5, 30, 1/3)
[1]  7 10 13 11  7
```

IL RISULTATO DELLA rbinom E' CHE SU 5 TENTATIVI, TIZIO PRENDERA' UN 7, UN 10, UN 13, UN 11 E UN 7. IL RISULTATO CAMBIA OGNI VOLTA CHE SI FA GIRARE LA rbinom!

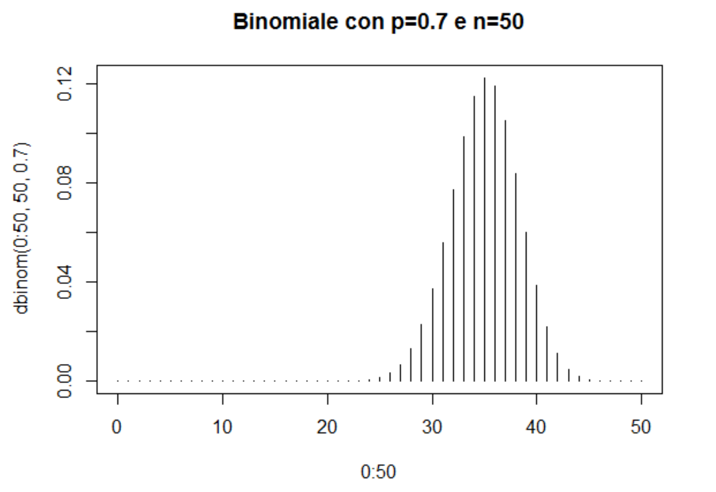
```
> plot(0:5, dbinom(0:5, 5, 0.7), "h")
> title(main="Binomiale con p=0.7 e n=5")
```



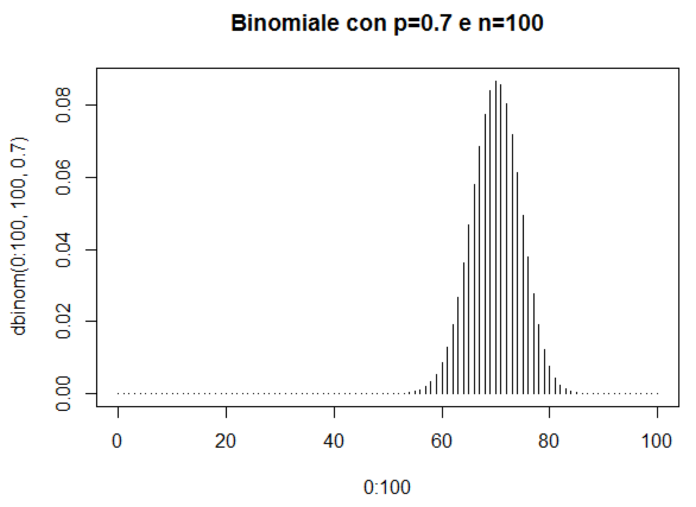

```
> plot(0:20, dbinom(0:20, 20, 0.7), "h")
> title(main="Binomiale con p=0.7 e n=20")
```




```
> plot(0:50, dbinom(0:50, 50, 0.7), "h")
> title(main="Binomiale con p=0.7 e n=50")
```




```
> plot(0:100, dbinom(0:100, 100, 0.7), "h")
> title(main="Binomiale con p=0.7 e n=100")
```



LA VARIABILE DI POISSON (EVENTI RARI)

La variabile descrive quante volte si presenta un evento aleatorio, di probabilità infinitesima nella singola prova ($p < 0{,}1$), in un dato intervallo di tempo.

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

λ = n. medio di volte in cui si verifica l'evento

LA VARIABILE DI POISSON

I momenti della variabile di Poisson sono:

- Media: $\mu = \lambda$
- Varianza: $\sigma^2 = \lambda$
- Scarto quadratico medio: $\sigma = \sqrt{\lambda}$

LA VARIABILE DI POISSON

In R si definiscono quattro funzioni per la variabile di Poisson:

- **dpois()** calcola la densità di probabilità
- **ppois()** è la funzione di probabilità cumulata
- **qpois()** è l'inversa della probabilità cumulata
- **rpois()** per creare dei valori random generati da una variabile casuale di Poisson

ESEMPIO DI VARIABILE DI POISSON

La probabilità che un macchinario, che produce migliaia di pezzi, ne produca uno difettoso in un'ora è in media molto bassa e pari a $\lambda=2$.

Descrivere con una opportuna variabile casuale la probabilità di avere un numero di pezzi difettosi all'ora da 0 a 5.

ESEMPIO DI VARIABILE DI POISSON

λ = 2

k = 0, 1, 2, 3, 4, 5, …

$$\frac{\lambda^k}{k!} e^{-\lambda}$$

ESEMPIO DI VARIABILE DI POISSON

k	P(k)
0	0,135335
1	0,270671
2	0,270671
3	0,180447
4	0,090224
5	0,036089
6 e oltre	0,016563

LA FUNZIONE dpois(k, λ)

 # CREO IL VETTORE DEI k

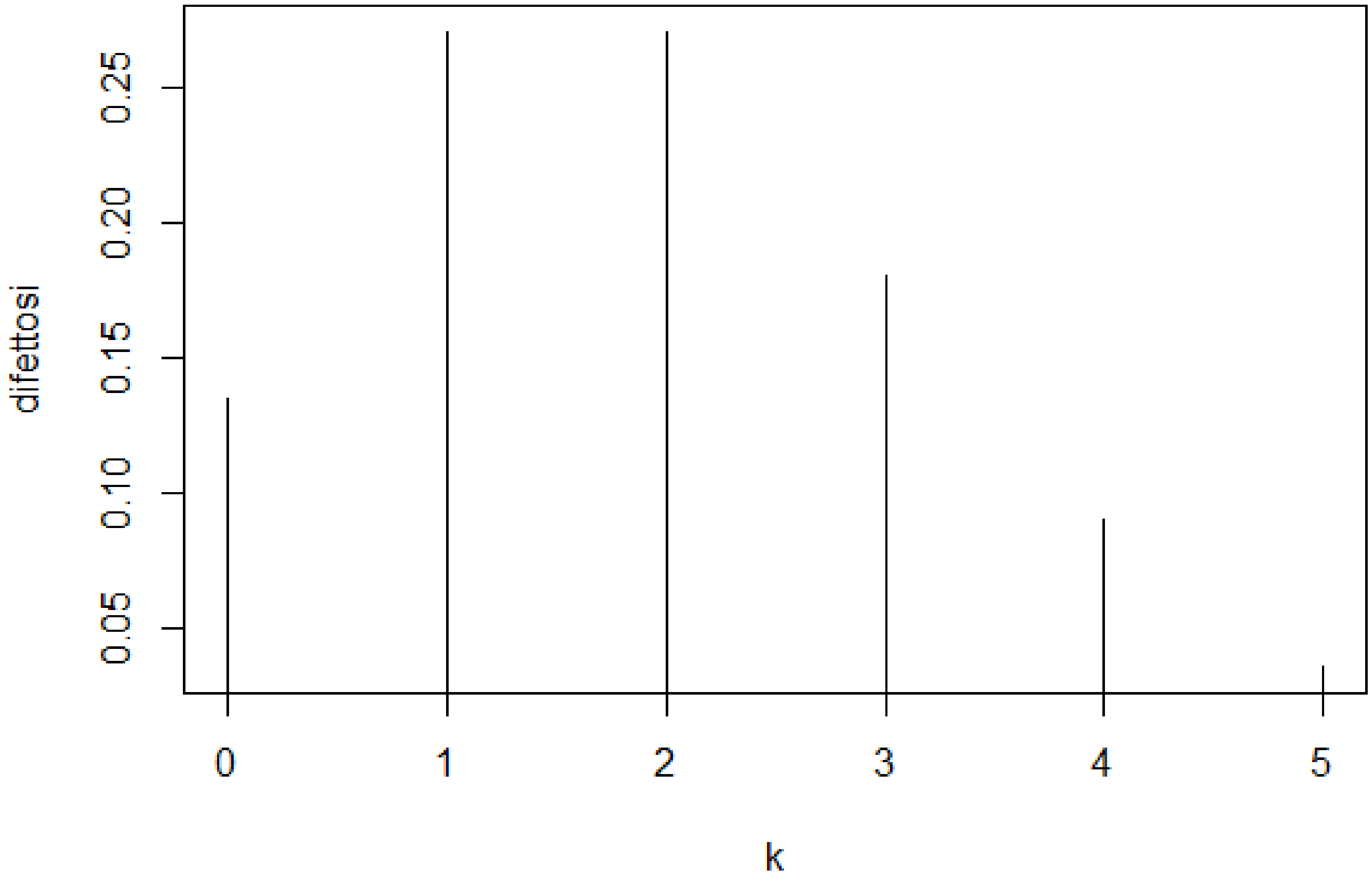
```
> k=c(0:5)
> k
[1] 0 1 2 3 4 5
```

 # CALCOLO LE PROBABILITA’ DELLA POISSON CON LA FUNZIONE dpois

```
> difettosi=dpois(k, 2) # dpois(k, λ)
> difettosi
[1] 0.13533528 0.27067057 0.27067057 0.18044704
0.09022352 0.03608941
```

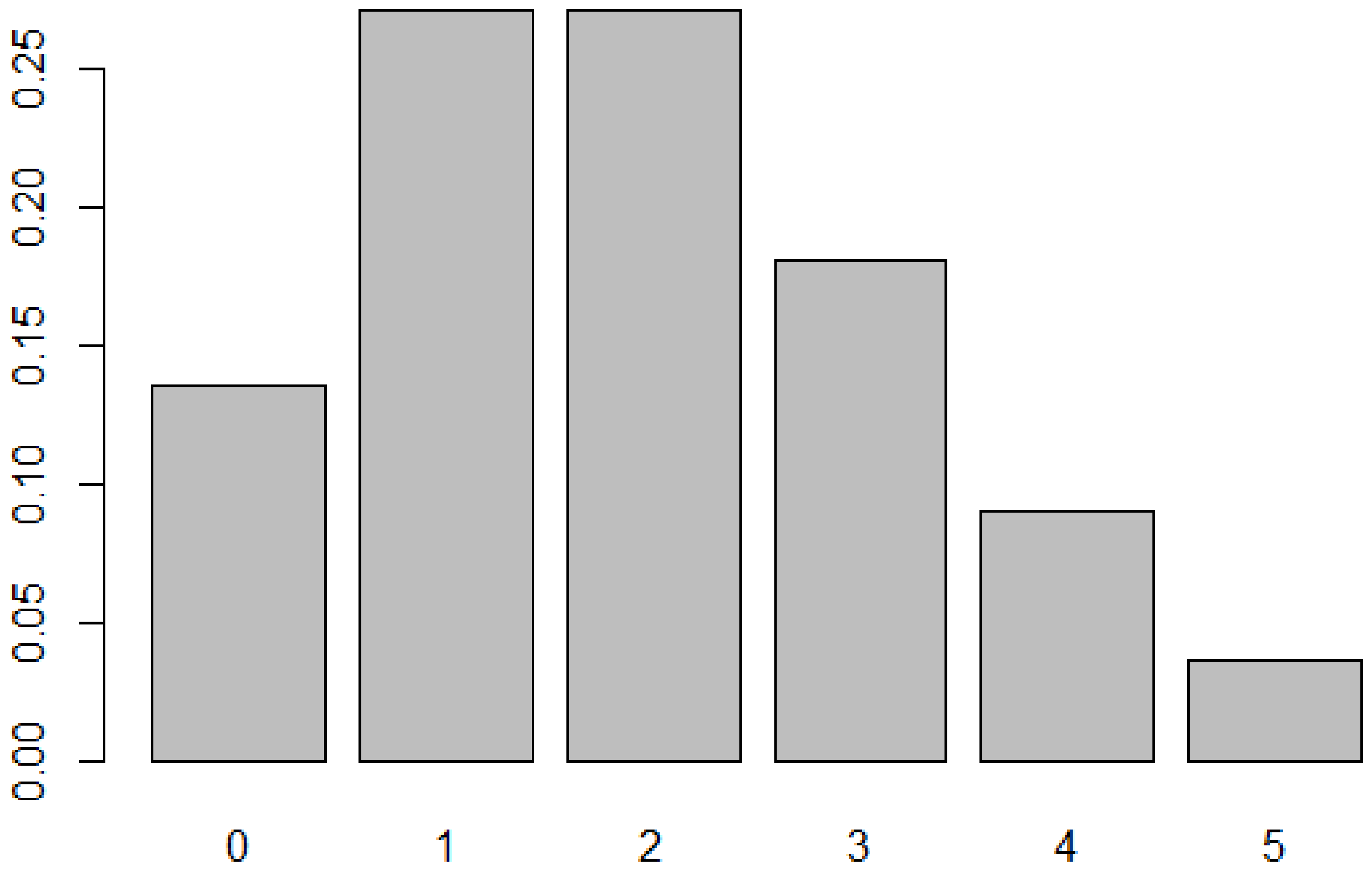
DISEGNO IL GRAFICO DELLA VARIABILE

> plot(k, difettosi, "h") **# 'h' CREA LE LINEE**



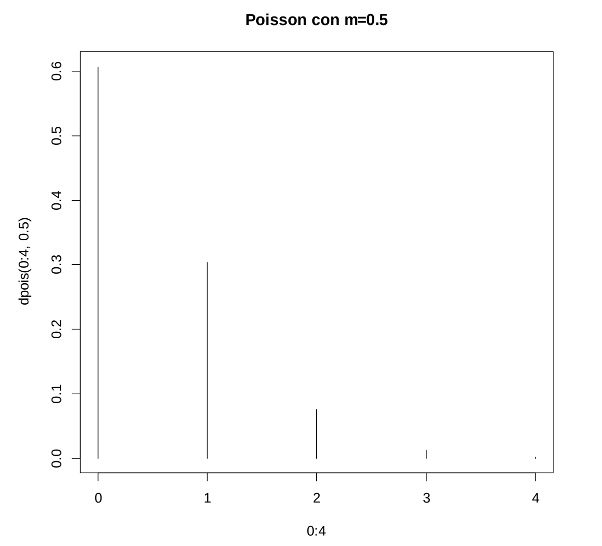
PER UN GRAFICO PIU' ELEGANTE

```
> barplot(difettosi, names.arg=k)
```



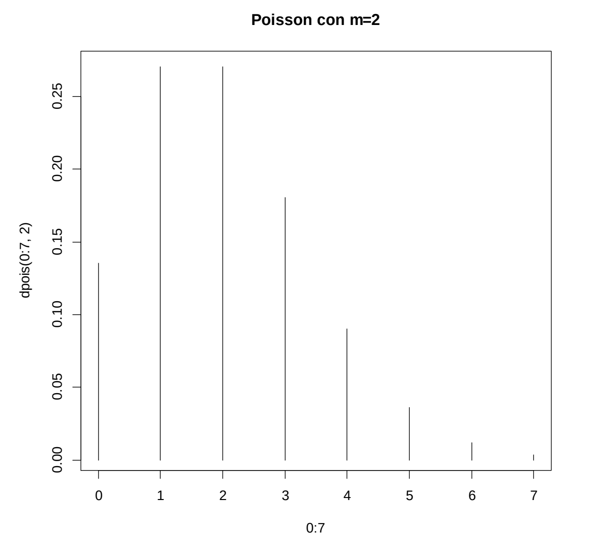
```
> plot(0:4, dpois(0:4, 0.5), "h")

> title(main="Poisson con m=0.5")
```



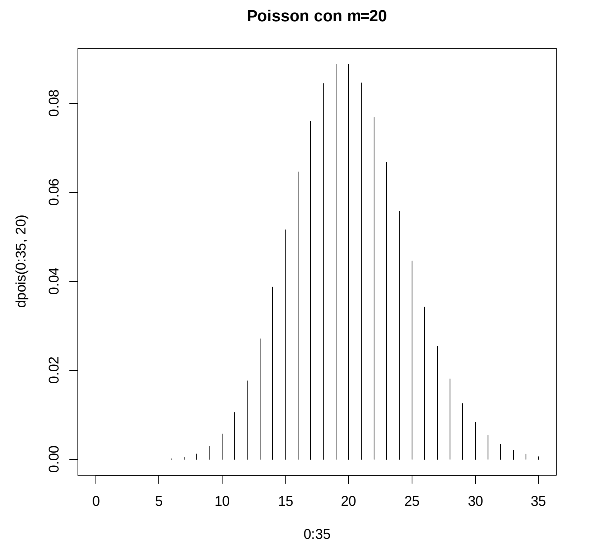

```
> plot(0:7, dpois(0:7, 2), "h")

> title(main="Poisson con m=2")
```




```
> plot(0:35, dpois(0:35, 20), "h")

> title(main="Poisson con m=20")
```




```
> plot(0:150, dpois(0:150, 100), "h")

> title(main="Poisson con m=100")
```