Esempio.

La prevalenza del daltonismo nella popolazione maschile è $p = 6\%$.

Qual è la probabilità di avere almeno 2 daltonici in un campione di 25?

Il numero di daltonici in una popolazione di 25 è una v.a. $B(25, p)$

La probabilità di avere esattamente k daltonici è $\binom{25}{k} p^k (1-p)^{25-k}$.

Quindi la risposta è $\sum_{k=2}^{25} \binom{25}{k} p^k (1-p)^{25-k} = 45\%$

```
sum( dbinom( 2:25, 25, 0.06 ) )
```

Un'altra risposta è $1 - p'$ dove p' è la probabilità di avere < 2 daltonici.

Poiché $p' = \binom{25}{0} p^0 (1-p)^{25} + \binom{25}{1} p^1 (1-p)^{25-1}$

Otteniamo $1 - (1-p)^{25} - 25\,p\,(1-p)^{24}$. (Più semplice da calcolare)

Per $i = 1, 2$ siano X_i variabili aleatorie indipendenti $B(1, p)$.

Esempio: $\Omega = \boldsymbol{U}^2$, ovvero due estrazioni dall'urna di biglie rosse e blu.

X_1 è 0 se la 1^{a} biglia estratta è rossa, 1 se è blu.

X_2 è 0 se la 2^{a} biglia estratta è rossa, 1 se è blu.

$$X := X_1 + X_2 \; = \begin{cases} 0 & \text{se nessuna biglia estratta è blu} \\ 1 & \text{se solo una delle due biglia estratte è blu} \\ 2 & \text{se entrambe le biglie estratte sono blu} \end{cases}$$

Quindi X è una variabile aleatoria $B(2, p)$.

Possiamo verificare che:

$P(X{=}0) = (1-p)^2 \qquad P(X{=}1) = 2p(p-1) \qquad P(X{=}2) = p^2$

Somma di v.a.i. $B(n, p)$.

In generale:

- Se $i = 1, 2$ sono X_i v.a.i. con distribuzione $B(n, p)$ e $B(m, p)$.

  Allora $X := X_1 + X_2$ è una variabile aleatoria $B(n + m, p)$.

- E se $i = 1, \ldots, n$ sono X_i v.a.i. con distribuzione $B(1, p)$.

  Allora $X := X_1 + \cdots + X_n$ è una variabile aleatoria $B(n, p)$.

Esempio (test per dadi difettosi).

Una scatola di dadi contiene dei dadi difettosi che producono ⚂ con probabilità $> 1/6$. Indichiamo con D l'insieme dei dadi difettosi.

Progettiamo un test per individuare i dadi difettosi. Lanciamo un dado n volte, se ⚂ ha frequenza $\geq 1/6 + \varepsilon$ diciamo che il test è **positivo**. Vogliamo valori ragionevoli di n ed ε per concludere che il dado è difettoso. (Scriviamo X per la v.a. che conta quante volte esce ⚂.)

- Nella terminologia dei test diagnostici la **specificità** del test è

$$P(-|\neg D) \ = \ P\left(\frac{1}{n}X < \frac{1}{6} + \varepsilon\right) \ = \sum_{k/n < 1/6+\varepsilon} \binom{n}{k} \left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{n-k}.$$

- Nei test statistici si preferisce parlare **livello di significatività** del test che è la probabilità di un falso positivo (1−specificità):

$$\alpha = P(+|\neg D) = 1 - P(-|\neg D).$$

  In statistica i falsi positivi si chiamano **errori del primo tipo**. L'affermazione "*dado* $\in \neg D$" si chiama **ipotesi nulla** e si denota con $\mathbf{H_0}$. Quando il test è positivo si dice $\mathbf{H_0}$ **viene rigettata**.

Esempio (test per dadi difettosi).

Abbiamo ristretto i possibili difetti ad uno preciso: ⚂ ha probabilità $> 1/6$. Questa descrizione del difetto si chiama **ipotesi alternativa** e viene spesso denotata da $\boldsymbol{H_A}$. Quindi H_0 e H_A assieme restringono il campo delle possiblità (qui dichiariamo il caso in cui la probabilità è $< 1/6$ impossibile o irrilevante).

Per discutere la sensibilità del test è anche necessario concordare su cosa intendiamo per difettoso.

Consideriamo un ben preciso dado d. Anche se la vera probabilità del ⚂ non è misurabile, assumiamo (un esperimento mentale) di conoscerla con esattezza. Supponiamo sia $1/6 + 10^{-27}$. Poiché 10^{-27} è un numero spaventosamente piccolo, potremmo anche concordare che d non sia difettoso.

Serve fissare la **differenza** o **effect size**, che qui noi indicheremo con δ. I dadi che lanciano ⚂ con probabilità $\leq 1/6 + \delta$ il dado verrà considerato come non difettoso.

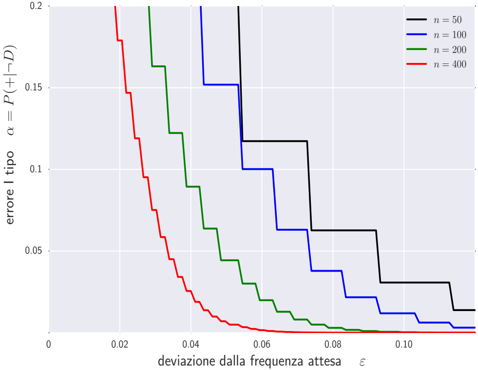
Esempio (test per dadi difettosi).

- Nella terminologia dei test diagnostici: la **sensibilità** del test è

$$\begin{aligned} P(+|D) &= P\left(\frac{1}{n}X \geq 1/6 + \varepsilon\right) \\ &= \sum_{k/n \geq 1/6+\varepsilon} \binom{n}{k} \left(\frac{1}{6} + \delta\right)^k \left(1 - \frac{1}{6} - \delta\right)^{n-k} \end{aligned}$$

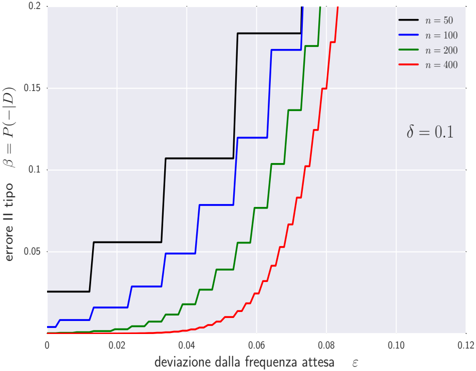
- Nella teminologia dei test statistici la sensibilità si chiama **potenza del test**. La probabilità di commettere un errore del II tipo si indica generalmente con $\beta = P(-|D)$. La potenza del test è quindi $1 - \beta$.

Esempio (test per dadi difettosi).



$$\alpha \;=\; P(+|\neg D) \;=\; \sum_{k/n \geq 1/6+\varepsilon} P(X{=}k)$$

$$X \sim \mathrm{B}(n, \frac{1}{6})$$



$$\beta \;=\; P(-|D) \;=\; \sum_{k/n<1/6+\varepsilon} P(X{=}k)$$

$$X \sim \mathrm{B}(n, \frac{1}{6} + \delta)$$

Esempio (test per dadi difettosi).

In molti casi non abbiamo la possibilità di progettare un esperimento in anticipo. Per esempio, di un singolo dado ci viene detto che su 76 lanci il ⚂ è uscito 18 volte. Dobbiamo decidere se il dado è equilibrato o no.

Il minimo ε che ci permette di considerare il risultato positivo (ovvero rifiuare H_0) è $\varepsilon = 18/76 - 1/6 = 0.083$. Con $\varepsilon = 0.083$ e $n = 76$ la probabità di falso positivo positivo (errore del I tipo) è

$$P(+|\neg D) \;=\; P\big(X \geq 18\big) \;= \sum_{k \geq 18} \binom{76}{k} \left(\frac{1}{6}\right)^{k} \left(1 - \frac{1}{6}\right)^{76-k} \;=\; 7.3\%$$

A parole diremo che l'ipotesi nulla può essere rifiuata con p-valore 0.073.

Schema terminologia test statistici (e diagnostici)

$P(+\|\neg D)$ Prob. errore I tipo (falso pos.) Livello di significatività Tipicamente indicato con α	$P(+\|D)$ Potenza (Sensibilità) Tipicamente indicato con $1-\beta$
$P(-\|\neg D)$ Tipicamente indicato con $1-\alpha$ (Specificità)	$P(-\|D)$ Prob. errore II tipo (falso neg.) Tipicamente indicato con β

$+ =$ rifiutare (reject) H_0

$- =$ non rifiutare (reject) H_0