STATISTICA

VERIFICA D'IPOTESI - 2

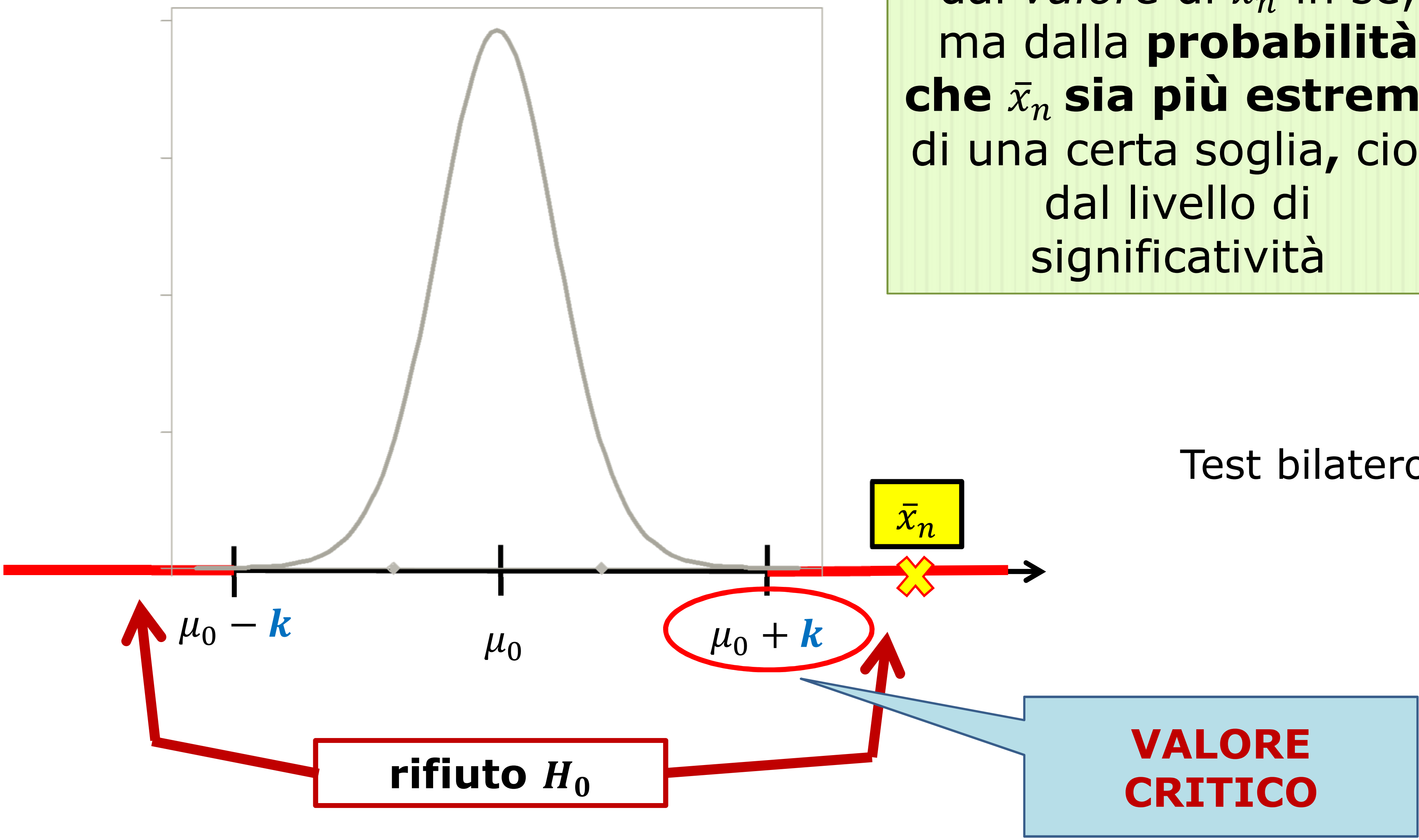
Verifica d'ipotesi: p

$(X_1, \dots, X_n)$ campione aleatorio $bern(p)$, $\boldsymbol{np_0 \geq 5} \;\&\; \boldsymbol{n(1-p_0) \geq 5}$

$\boldsymbol{H_0}$	$\boldsymbol{H_1}$	si rifiuta $\boldsymbol{H_0}$ se:
$p = p_0$	$p \neq p_0$	$\left\lvert \dfrac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} \right\rvert > z_{\alpha/2}$
$p = p_0$	$p > p_0$	$\dfrac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} > z_\alpha$
$p = p_0$	$p < p_0$	$\dfrac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} < -z_\alpha$

Il rifiuto non dipende dal *valore* di $\bar{x}_n$ in sè, ma dalla **probabilità che $\bar{x}_n$ sia più estremo** di una certa soglia**,** cioè dal livello di significatività

Test bilatero



Prob. di finire nella zona rossa = α, livello di significatività

Esercizio 4

Vogliamo verificare che una moneta sia effettivamente equilibrata. La lanciamo 500 volte ottenendo un numero di T pari a 274. Sottoporre a verifica l'ipotesi nulla che la moneta sia equilibrata, al livello dell'1%.

Esercizio 4

Vogliamo verificare che una moneta sia effettivamente equilibrata. La lanciamo 500 volte ottenendo un numero di T pari a 274. Sottoporre a verifica l'ipotesi nulla che la moneta sia equilibrata, al livello dell'1%.

$(X_1, \dots, X_{500})$, i.i.d $b(p)$ $\quad H_0 : p = p_0 = 0.5 \quad H_1 : p \neq 0.5$

$$\left|\frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}}\right| = \left|\frac{\frac{274}{500} - 0.5}{\sqrt{0.5 \times 0.5/500}}\right| = \left|\frac{0.548 - 0.5}{0.02236068}\right| = \frac{0.048}{0.02236068} = 2.10$$

Esercizio 4

Vogliamo verificare che una moneta sia effettivamente equilibrata. La lanciamo 500 volte ottenendo un numero di T pari a 274. Sottoporre a verifica l'ipotesi nulla che la moneta sia equilibrata, al livello dell'1%.

$(X_1, \ldots, X_{500})$, i.i.d $b(p)$ $\quad H_0 : p = p_0 = 0.5 \quad H_1 : p \neq 0.5$

$$\left| \frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} \right| = \left| \frac{\frac{274}{500} - 0.5}{\sqrt{0.5 \times 0.5/500}} \right| = \left| \frac{0.548 - 0.5}{0.02236068} \right| = \frac{0.048}{0.02236068} = 2.10$$

$$z_{0.01/2} = 2.5758$$

Esercizio 4

Vogliamo verificare che una moneta sia effettivamente equilibrata. La lanciamo 500 volte ottenendo un numero di T pari a 274. Sottoporre a verifica l'ipotesi nulla che la moneta sia equilibrata, al livello dell'1%.

$(X_1, \dots, X_{500})$, i.i.d $b(p)$ $\quad H_0 : p = p_0 = 0.5 \quad H_1 : p \neq 0.5$

$$\left|\frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}}\right| = \left|\frac{\frac{274}{500} - 0.5}{\sqrt{0.5 \times 0.5/500}}\right| = \left|\frac{0.548 - 0.5}{0.02236068}\right| = \frac{0.048}{0.02236068} = 2.10$$

$$z_{0.01/2} = 2.5758$$

$2.10 < 2.5758.96 \Longrightarrow$ **non rifiuto** al livello dell'1% l'ipotesi nulla

Esercizio 4

Vogliamo verificare che una moneta sia effettivamente equilibrata. La lanciamo 500 volte ottenendo un numero di T pari a 274. Sottoporre a verifica l'ipotesi nulla che la moneta sia equilibrata, al livello dell'1%.

$(X_1, \dots, X_{500})$, i.i.d $b(p)$ $\quad H_0 : p = p_0 = 0.5 \quad H_1 : p \neq 0.5$

$$\left|\frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}}\right| = \left|\frac{\frac{274}{500} - 0.5}{\sqrt{0.5 \times 0.5/500}}\right| = \left|\frac{0.548 - 0.5}{0.02236068}\right| = \frac{0.048}{0.02236068} = 2.10$$

$$z_{0.01/2} = 2.5758$$

$2.10 < 2.5758.96 \Longrightarrow$ **non rifiuto** al livello dell'1% l'ipotesi nulla

E al livello del 5% si potrebbe rifiutare?

Esercizio 5

La spesa media per alimentari e bevande non alcoliche in un campione di 1200 famiglie, nel 2016, è risultata di 445.10€ con una deviazione standard di 73.75 €.

a) Costruire un intervallo di confidenza del 95% per la spesa media mensile familiare per alimentari e bevande alcoliche nella popolazione di riferimento.

Esercizio 5

La spesa media per alimentari e bevande non alcoliche in un campione di 1200 famiglie, nel 2016, è risultata di 445.10€ con una deviazione standard di 73.75 €.

a) Costruire un intervallo di confidenza del 95% per la spesa media mensile familiare per alimentari e bevande alcoliche nella popolazione di riferimento.

$X_1, X_2, \dots, X_n$ i.i.d

$X_i \sim N(\mu,\, \sigma^2)$

o TLC

$$\left(\bar{X}_n - t(n-1)_{\alpha/2} \times \sqrt{\frac{{S_n}^2}{n}}, \bar{X}_n + t(n-1)_{\alpha/2} \times \sqrt{\frac{{S_n}^2}{n}}\right)$$

Esercizio 5

La spesa media per alimentari e bevande non alcoliche in un campione di 1200 famiglie, nel 2016, è risultata di 445.10€ con una deviazione standard di 73.75 €.

a) Costruire un intervallo di confidenza del 95% per la spesa media mensile familiare per alimentari e bevande alcoliche nella popolazione di riferimento.

$X_1, X_2, \dots, X_n$ i.i.d

$X_i \sim N(\mu, \sigma^2)$

o TLC

$$\left(\bar{X}_n - t(n-1)_{\alpha/2} \times \sqrt{\frac{{S_n}^2}{n}}, \bar{X}_n + t(n-1)_{\alpha/2} \times \sqrt{\frac{{S_n}^2}{n}}\right)$$

$$\left(\bar{x}_{1200} - z_{0.05/2} \times \frac{73.75}{\sqrt{1200}}, \bar{x}_{1200} + z_{0.05/2} \times \frac{73.75}{\sqrt{1200}}\right) =$$

$$\left(445.10 - 1.96 \times \frac{73.75}{\sqrt{1200}}, 445.10 + 1.96 \times \frac{73.75}{\sqrt{1200}}\right) = (440.93, 449.27)$$

Esercizio 5

La spesa media per alimentari e bevande non alcoliche in un campione di 1200 famiglie, nel 2016, è risultata di 445.10€ con una deviazione standard di 73.75 €.

b) Secondo gli ultimi dati ISTAT, del 2014, la spesa media mensile di una famiglia italiana per alimentari e bevande non alcoliche è stata di 436.06€ (*). Il campione contiene abbastanza evidenza per poter affermare, ad un livello di significatività del 5%, che la spesa familiare media mensile delle famiglie italiane per questo settore è aumentata?

(*) ISTAT- La spesa per i consumi delle famiglie. Comunicato 8 luglio 2015.

Esercizio 5

La spesa media per alimentari e bevande non alcoliche in un campione di 1200 famiglie, nel 2016, è risultata di 445.10€ con una deviazione standard di 73.75 €.

b) Secondo gli ultimi dati ISTAT, del 2014, la spesa media mensile di una famiglia italiana per alimentari e bevande non alcoliche è stata di 436.06€. Il campione contiene abbastanza evidenza per poter affermare, ad un livello di significatività del 5%, che la spesa familiare media mensile delle famiglie italiane per questo settore è aumentata?

$X_1, X_2, \dots, X_n$ i.i.d

$X_i \sim N(\mu, \sigma^2)$

o TLC

$H_0: \mu = \mu_0 = 436.06 \quad H_1: \mu > 436.06$

$$\frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t(n-1)_\alpha$$

Esercizio 5

La spesa media per alimentari e bevande non alcoliche in un campione di 1200 famiglie, nel 2016, è risultata di 445.10€ con una deviazione standard di 73.75 €.

b) Secondo gli ultimi dati ISTAT, del 2014, la spesa media mensile di una famiglia italiana per alimentari e bevande non alcoliche è stata di 436.06€. Il campione contiene abbastanza evidenza per poter affermare, ad un livello di significatività del 5%, che la spesa familiare media mensile delle famiglie italiane per questo settore è aumentata?

$X_1, X_2, \dots, X_n$ i.i.d

$H_0 : \mu = \mu_0 = 436.06 \quad H_1 : \mu > 436.06$

$X_i \sim N(\mu, \sigma^2)$

o TLC

$$\boxed{\frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t(n-1)_\alpha}$$

$(n = 1200)$

$$\frac{445.10 - 436.06}{73.75/\sqrt{1200}} = 4.25$$

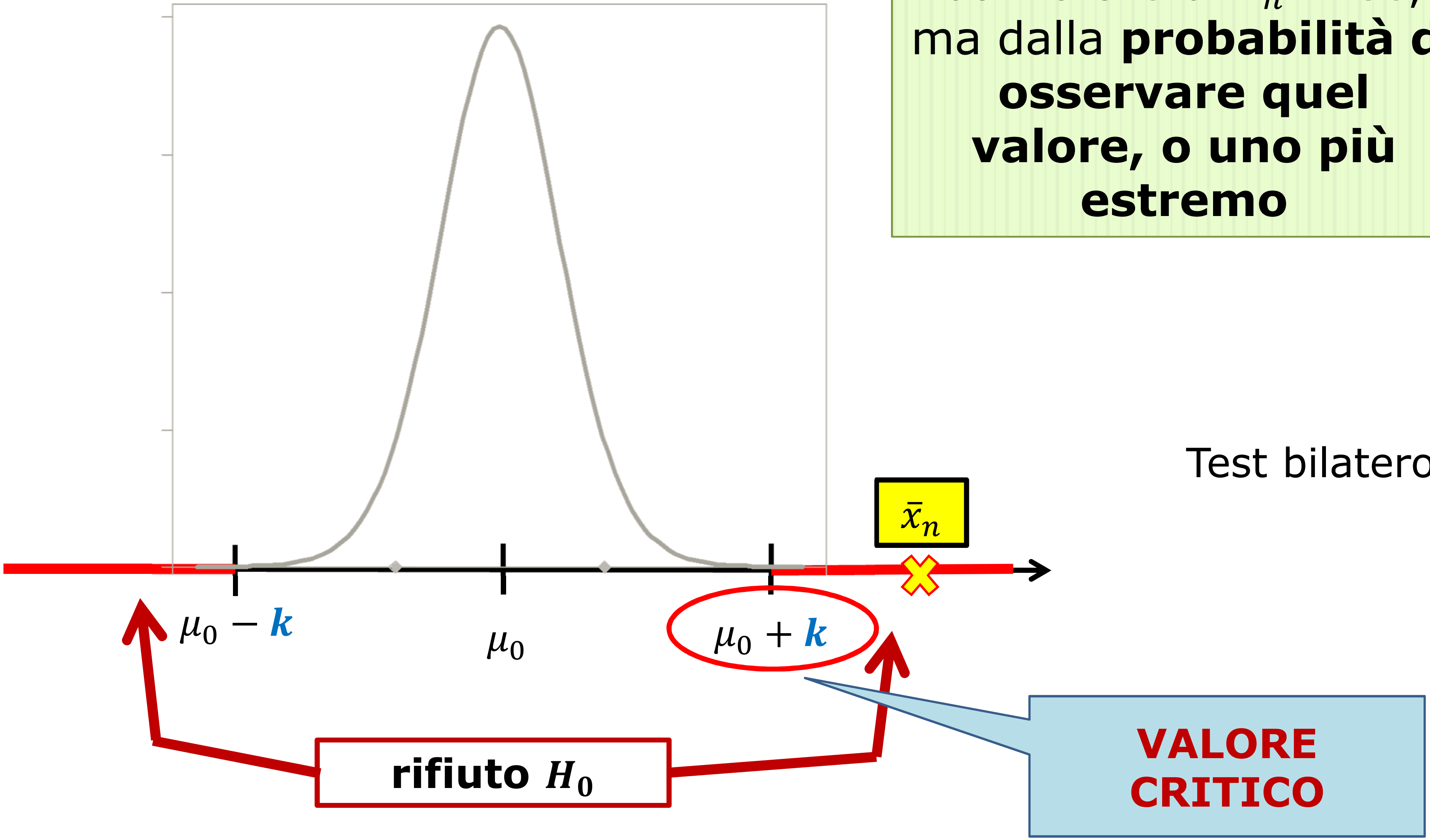
$$= z_{0.05} = 1.645$$

SI', al livello del 5%

Il *p-value*

Il rifiuto non dipende dal *valore* di $\bar{x}_n$ in sè, ma dalla **probabilità di osservare quel valore, o uno più estremo**

Test bilatero



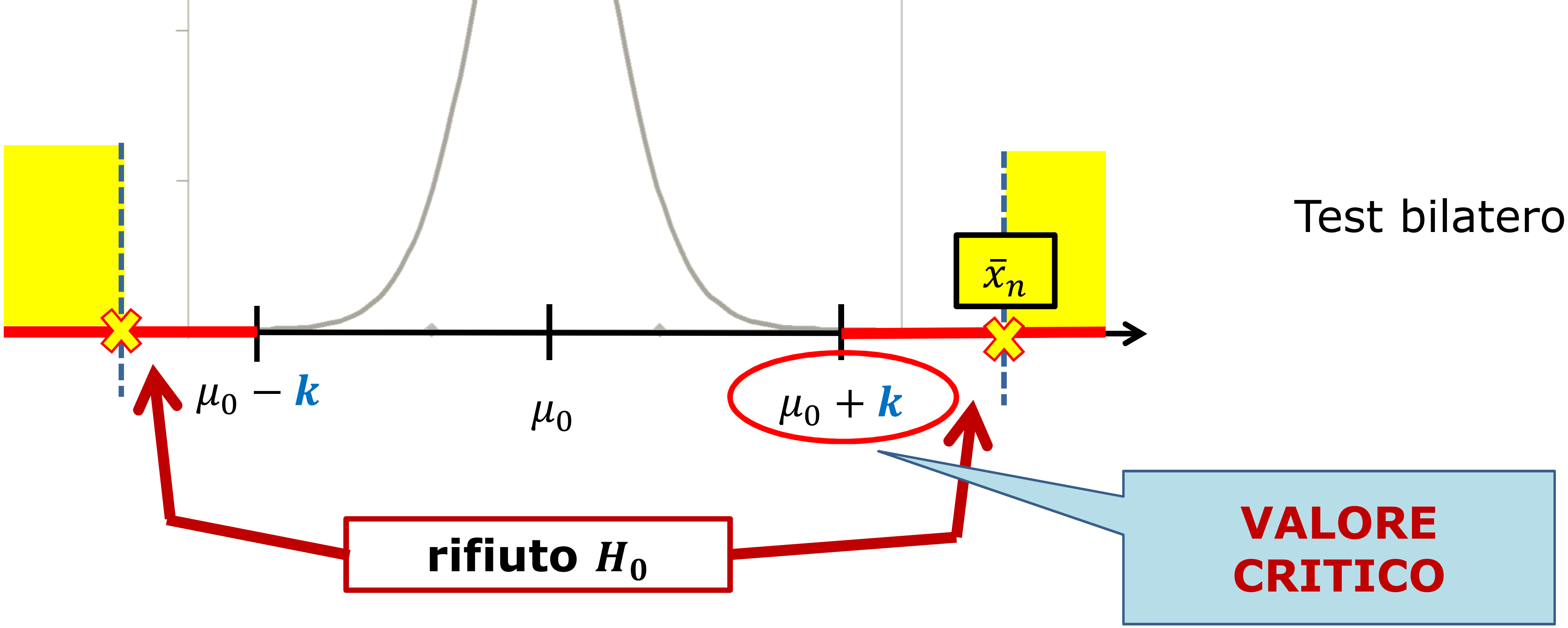
Prob. di finire nella zona rossa = $\boldsymbol{\alpha}$, livello di significatività

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

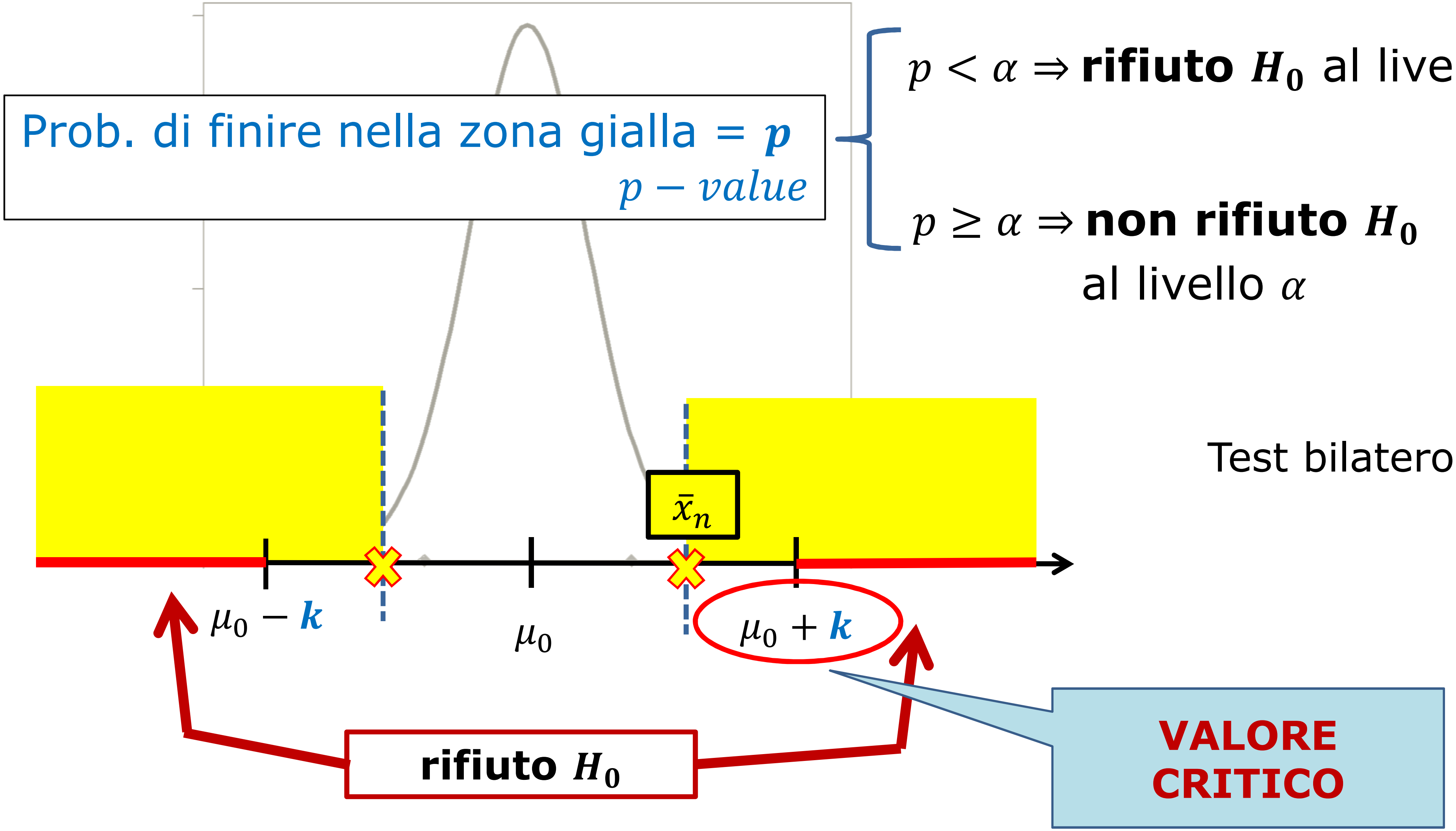
$p - value$

$p < \alpha \Rightarrow$ **rifiuto $\boldsymbol{H_0}$** al livello α



Prob. di finire nella zona rossa = $\boldsymbol{\alpha}$, livello di significatività

Il *p-value*



$p < \alpha \Rightarrow$ **rifiuto $\boldsymbol{H_0}$** al livello α

$p \geq \alpha \Rightarrow$ **non rifiuto $\boldsymbol{H_0}$** al livello α

Test bilatero

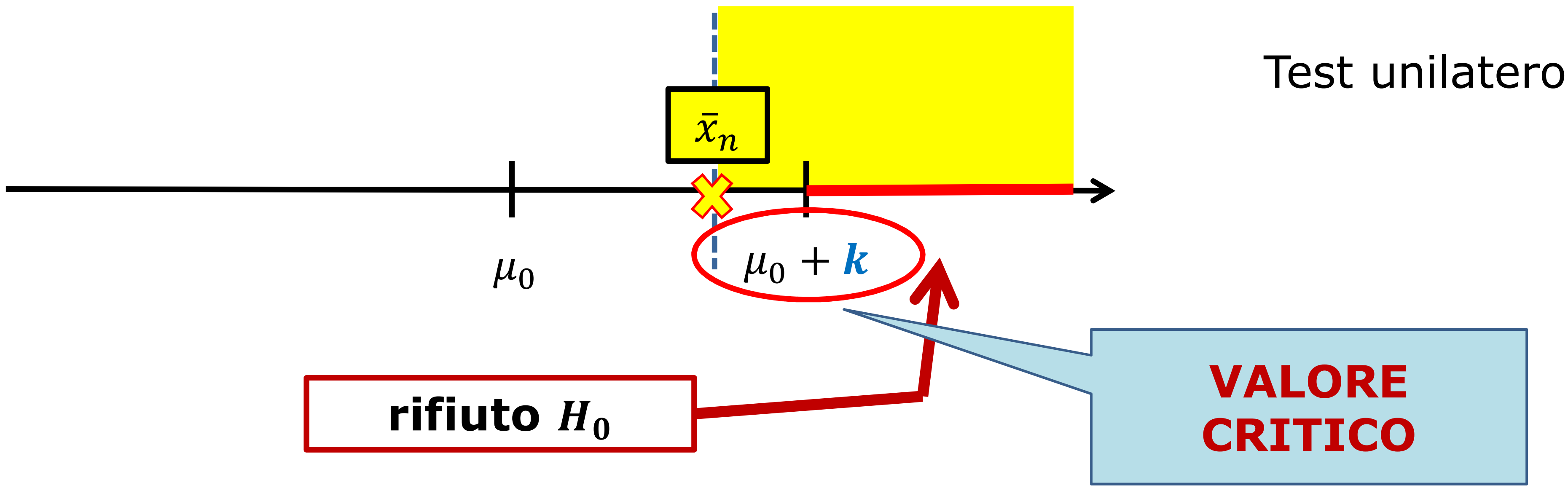
Prob. di finire nella zona rossa = $\boldsymbol{\alpha}$, livello di significatività

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$
$p - value$

$p < \alpha \Rightarrow$ **rifiuto** $\boldsymbol{H_0}$ al livello α

$p \geq \alpha \Rightarrow$ **non rifiuto** $\boldsymbol{H_0}$ al livello α



Prob. di finire nella zona rossa = $\boldsymbol{\alpha}$, livello di significatività

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

$p - value$

$$\begin{cases} p < \alpha \Rightarrow \textbf{rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \\ p \geq \alpha \Rightarrow \textbf{non rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \end{cases}$$

$p \in (0,1)$

"… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age (p = 0.98) was not statistically significant."

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

$p - value$

$$\begin{cases} p < \alpha \Rightarrow \textbf{rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \\ p \geq \alpha \Rightarrow \textbf{non rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \end{cases}$$

$p \in (0,1)$

"… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age 98) was not statistically significant."

è stato fatto un test, $p < \alpha = 0.05 \Rightarrow$ si rifiuta l'ipotesi nulla (che genere del paziente e sito del melanoma non siano fattori rilevanti) al livello $\alpha = 0.05$

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$ $\begin{cases} p < \alpha \Rightarrow \textbf{rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \\ p \geq \alpha \Rightarrow \textbf{non rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \end{cases}$

$p - value$

$p \in (0,1)$

"… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age (p = 0.98) was not statistically significant."

è stato fatto un test, $p > \alpha = 0.05 \Rightarrow$ non si può rifiutare l'ipotesi nulla (che l'età del paziente non sia un fattore rilevante) al livello $\alpha = 0.05$

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

$p - value$

$$p \in (0,1)$$

$p < \alpha \Rightarrow$ **rifiuto** $\boldsymbol{H_0}$ al livello α

$p \geq \alpha \Rightarrow$ **non rifiuto** $\boldsymbol{H_0}$ al livello α

"… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age [illegible]98) was not statistically significant."

è stato fatto un test, $p < \boldsymbol{\alpha = 0.01} \Rightarrow$ si rifiuta l'ipotesi nulla (che genere del paziente e sito del melanoma non siano fattori rilevanti) al livello $\boldsymbol{\alpha = 0.01}$

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

$p - value$

$$\begin{cases} p < \alpha \Rightarrow \textbf{rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \\ p \geq \alpha \Rightarrow \textbf{non rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \end{cases}$$

$p \in (0,1)$

> "… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age [illegible]98) was not statistically significant."

è stato fatto un test, $p < \boldsymbol{\alpha = 0.001} \Rightarrow$????????

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

$p - value$

$$\begin{cases} p < \alpha \Rightarrow \textbf{rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \\ p \geq \alpha \Rightarrow \textbf{non rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \end{cases}$$

$p \in (0,1)$

"… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age 98) was not statistically significant."

è stato fatto un test, $p < \boldsymbol{\alpha = 0.001} \Rightarrow$????????

Si rifiuta l'ipotesi che il genere del paziente sia irrilevante al livello $\boldsymbol{\alpha = 0.001}$

NON si rifiuta l'ipotesi che il sito del melanoma sia irrilevante al livello $\boldsymbol{\alpha = 0.001}$ $(0.002 \geq 0.001)$

Il *p-value*

Prob. di finire nella zona gialla = $\boldsymbol{p}$

$p - value$

$$\begin{cases} p < \alpha \Rightarrow \textbf{rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \\ p \geq \alpha \Rightarrow \textbf{non rifiuto } \boldsymbol{H_0} \text{ al livello } \alpha \end{cases}$$

$p \in (0,1)$

"… the sex of patients and site of melanoma also were statistically significant (p = 0.00001 and 0.002 respectively), whereas age 98) was not statistically significant."

Più il $p - value$ è piccolo e più è forte la significatività del test, cioè l'evidenza **contro $\boldsymbol{H_0}$**.

è stato fatto un test, $p < \boldsymbol{\alpha} =$

Si rifiuta l'ipotesi che il genere del paziente sia irrilevante al livello $\boldsymbol{\alpha = 0.001}$

NON si rifiuta l'ipotesi che il sito del melanoma sia irrilevante al livello $\boldsymbol{\alpha = 0.001}$ $(0.002 \geq 0.001)$

Esercizio 7

“Swiss-wide the month May to October show significantly increasing trends of rainfall erosivity for the observed period ($p<0.005$). Only in February a significantly decreasing trend of rainfall erosivity is found ($p<0.01$).”

Esercizio 7

"Swiss-wide the month May to October show significantly increasing trends of rainfall erosivity for the observed period (p<0.005). Only in February a significantly decreasing trend of rainfall erosivity is found (p<0.01)."

H_0 : assenza di trend crescente

H_1 : presenza di trend crescente

Rifiutiamo H_0 al livello: $\alpha = 0.05$?

$\alpha = 0.01$?

$\alpha = 0.001$?

Esercizio 7

"Swiss-wide the month May to October show significantly increasing trends of rainfall erosivity for the observed period (p<0.005). Only in February a significantly decreasing trend of rainfall erosivity is found (p<0.01)."

H_0 : assenza di trend decrescente

H_1 : presenza di trend decrescente

Rifiutiamo H_0 al livello: $\alpha = 0.05$?

$\alpha = 0.01$?

$\alpha = 0.001$?

Facciamo un salto indietro…

Era il 10/11 ….

Sottinteso c'è un test:
$H_0 : p = 0.04,\ H_1 : p \neq 0.04$

Es. 14 p. 136

In un test clinico sul Viagra è emerso che il 4% delle unità nel gruppo del placebo ha sofferto di mal di testa.

b) Applicando lo stesso tasso al gruppo del Viagra, determinare la prob. che su 8 soggetti che utilizzino il Viagra tutti lamentino il mal di testa.

$X \sim Bin(8, 0.04)$

$$P(X = 3) = \binom{8}{3} 0.04^3 (1 - 0.04)^5 = \frac{8 \times 7 \times 6}{3 \times 2} \times 0.000064 \times 0.815374 =$$
$$= 0.002922296 \approx 0.003$$

$$P(X = 8) = \binom{8}{8} 0.04^8 (1 - 0.04)^0 = 1 \times (6.55 \times 10^{-12}) \times 1 \approx 0$$

Facciamo un ...

Era il 10/11 ....

… e stiamo calcolando un *p-value*:
$P(X \geq 3) = 0.0031$
C'è abbastanza evidenza per rifiutare H_0 ?

Es. 14 p. 1...

In un test clinico sul Viagra è emerso che il 4% delle unità ... gruppo del placebo ha sofferto di mal di testa.

c) Se tutte le 8 unità che prendono il Viagra hanno mal di t... sembrerebbe che il tasso in questo gruppo sia diverso che quello del placebo…

$X \sim Bin(8, 0.04)$

$$P(X=3) = \binom{8}{3} 0.04^3 (1-0.04)^5 = \frac{8 \times 7 \times 6}{3 \times 2} \times 0.000064 \times 0.815374 =$$
$$= 0.002922296 \approx 0.003$$

$$P(X=8) = \binom{8}{8} 0.04^8 (1-0.04)^0 = 1 \times (6.55 \times 10^{-12}) \times 1 \approx 0$$

Se la prob. di avere il mal di testa col V fosse 0.04 (4%), sarebbe praticamente impossibile avere 8 su 8 col mal di testa! Ma **se li osservo, allora è l'ipotesi $p = 0.04$ che non è supportata dai dati.**

Definite il test sottinteso, indicate il *p-value* e concludete.

Esercizio 8

Era il 10/11 ….

Es. 6 pg. 139

Per i guidatori tra i 20 ed i 24 anni di età c'è un tasso annuo del 34% di incidenti d'auto (National Safety Council). Un investigatore assicurativo trova che in un gruppo di 500 guidatori scelti a caso a NY tra i 20 e i 24 anni il 42% ha avuto un incidente lo scorso anno.

c) Il risultato di NY sembra inusuale rispetto ai valori NSC? Si potrebbero giustificare assicurazioni più care per i Newyorkesi?

Tra i 500 di NY, $500 \times 0.42 = 210$ hanno avuto un incidente

$$X \sim Bin(500, 0.34) \Rightarrow P(149 \leq X \leq 191) = 0.96 \quad \text{(con software statistico)}$$

e $P(X \geq 210) = 0.00012$

Usiamo la "regola del *range*" per la distribuzione di X :

$(170 - 2 \times 10.59, 170 + 2 \times 10.59) = (\mathbf{148.82}, \mathbf{191.18}) \Rightarrow$ i 210 incidenti osservati a NY l'anno scorso costituiscono un valore inusuale, e potrebbero giustificare il più alto valore dei premi assicurativi.

STATISTICA

VERIFICA D'IPOTESI - 3

Due campioni **indipendenti**

Confrontiamo la spesa media annua per riparazioni dell'auto tra uomini e donne

Confrontiamo la longevità tra una popolazione isolana e una continentale

Confrontiamo l'effetto di un farmaco tra il gruppo del trattamento e quello del placebo

Confrontiamo il diametro delle uova deposte dai cuculi di Darwin in nidi di scricciolo o in nidi di pettirosso

Confrontiamo l'insorgere di una certa malattia tra chi fa sport e chi fa vita sedentaria

Due campioni **indipendenti**

Confrontiamo la spesa media annua per riparazioni dell'auto tra uomini e donne

Confrontiamo la longevità tra una popolazione isolana e una continentale

Confrontiamo l'effetto di un farmaco tra il gruppo del trattamento e quello del placebo

Confrontiamo il diametro delle uova deposte dai cuculi di Darwin in nidi di scricciolo o in nidi di pettirosso

Confrontiamo l'insorgere di una certa malattia tra chi fa sport e chi fa vita sedentaria

due gruppi di dati da due campioni **indipendenti**

Esempio 1

	Infarto miocardico			$\hat{p}_i$
Gruppo	**Sì**	**No**		
1-Placebo	189	10845	11034	0.0171
2-Aspirina	104	10933	11037	0.0094

$$H_0 : p_1 = p_2 \quad , \quad \boldsymbol{H_1 : p_1 > p_2}$$

Verifica d'ipotesi: p_1 e p_2

$(X_1, \dots, X_{n_1})$ campione aleatorio $bern(p_1)$ $\quad (n_1 p_1 \geq 5\ ecc.)$

$(Y_1, \dots, Y_{n_2})$ campione aleatorio $bern(p_2)$ $\quad (n_2 p_2 \geq 5\ ecc.)$

$\boldsymbol{H_0}$	$\boldsymbol{H_1}$	Rifiutiamo $\boldsymbol{H_0}$ se:
$p_1 = p_2$	$p_1 \neq p_2$	$\left\lvert \dfrac{\boldsymbol{\hat{p}_1 - \hat{p}_2}}{\sqrt{una\ varianza\ opportuna}} \right\rvert > z_{\alpha/2}$
$p_1 = p_2$	$p_1 > p_2$	$\dfrac{\boldsymbol{\hat{p}_1 - \hat{p}_2}}{\sqrt{una\ varianza\ opportuna}} > z_\alpha$
…	…	…

Verifica d'ipotesi: p_1 e p_2

$(X_1, \dots, X_{n_1})$ campione aleatorio $bern(p_1)$ $\quad (n_1 p_1 \geq 5\ ecc.)$

$(Y_1, \dots, Y_{n_2})$ campione aleatorio $bern(p_2)$ $\quad (n_2 p_2 \geq 5\ ecc.)$

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

stima della proporzione totale di "successi"

$$una\ varianza\ opportuna\ :\ \bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

Verifica d'ipotesi: p_1 e p_2

$(X_1, \dots, X_{n_1})$ campione aleatorio $bern(p_1)$ $\quad (n_1 p_1 \geq 5\ ecc.)$

$(Y_1, \dots, Y_{n_2})$ campione aleatorio $bern(p_2)$ $\quad (n_2 p_2 \geq 5\ ecc.)$

H_0	H_1	Rifiutiamo H_0 se:
$p_1 = p_2$	$p_1 \neq p_2$	$\left\lvert \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \right\rvert > z_{\alpha/2}$
$p_1 = p_2$	$p_1 > p_2$	$\dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} > z_{\alpha}$
…	…	…

Esempio 1

	Infarto miocardico			$\hat{p}_i$
Gruppo	**Sì**	**No**		
1-Placebo	189	10845	11034	0.0171
2-Aspirina	104	10933	11037	0.0094

$$H_0 : p_1 = p_2 \quad , \quad \boldsymbol{H_1 : p_1 > p_2}$$

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{189 + 104}{11034 + 11037} = 0.0133$$

$$\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = 0.0133 \times 0.9867 \times \left(\frac{1}{11034} + \frac{1}{11037}\right) = 0.000002378$$

Esempio 1

	Infarto miocardico			$\hat{p}_i$
Gruppo	**Sì**	**No**		
1-Placebo	189	10845	11034	0.0171
2-Aspirina	104	10933	11037	0.0094

$$H_0 : p_1 = p_2 \quad , \quad \boldsymbol{H_1 : p_1 > p_2}$$

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{189 + 104}{11034 + 11037} = 0.0133$$

$$\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = 0.0133 \times 0.9867 \times \left(\frac{1}{11034} + \frac{1}{11037}\right) = 0.000002378$$

$$\frac{0.0171 - 0.0094}{\sqrt{0.0000002378}} = 4.99 > z_{0.01} = 2.3663 \Rightarrow \text{c'è forte evidenza contro } H_0$$