Per il suo compleanno, il goloso Re di un lontano regno riceve in regalo da un altro sovrano un grande canestro contenente **4367** caramelle di tanti colori, tra cui **382** rosse.
Qualche tempo dopo il donatore impara che il Re ha già consumato **2093** caramelle, ma nessuna di quelle rosse. Rattristato di ciò, scrive al Re chiedendogli perché non ha gradito le caramelle rosse; ma questi gli risponde che non è così, egli ha scelto le caramelle senza fare caso al colore.
La risposta non convince il donatore…

DUE argomenti, entrambi necessari, inducono il donatore a dubitare della parola del Re:

- **un sospetto *"a priori"***: il donatore ritiene, prima di mettere alla prova il Re, che egli non ami particolarmente le caramelle di colore rosso.
- **la verifica statistica**: l'assenza di caramelle rosse in un grande numero di prelievi è assai poco probabile se la scelta non è "pilotata", e avvalora il sospetto dell'esistenza di un pregiudizio.

È vero o no che…

1) “12 è un numero pari”
2) “$\sqrt{2}$ non è un numero razionale”
3) “Le auto con targa pari consumano di più di quelle con targa dispari”
4) “Le auto con cambio automatico consumano di più di quelle con cambio manuale”
5) “L’aspirina fa passare il mal di testa”

6) “Questa moneta è equilibrata”.

TEOREMA: **$\sqrt{2}$ non è un numero razionale**

Dimostrazione **per assurdo**:

supponiamo che $\sqrt{2}$ sia razionale:

$$\sqrt{2}=\frac{m}{n}\,;\quad m^2=2n^2\,;\ m,n \text{ primi tra loro}$$

m è pari; quindi n è dispari (m,n sono primi tra loro)

Allora m^2 è multiplo di 4, mentre $2n^2$ non lo è

Conclusione:

IMPOSSIBILE!

La supposizione che $\sqrt{2}$ sia razionale è da respingere; dunque $\sqrt{2}$ è irrazionale.

Test sulla moneta.

Supponiamo che la moneta sia equilibrata

Il test: lanciamo per 200 volte la moneta

Risultato: 37 uscite di “testa”, 163 “croce”

Conclusione:

?

Sembra **poco probabile** che la moneta sia equilibrata, ma ***non saremo mai certi*** di ciò, né del suo contrario.

Il **rifiuto** o l'**accettazione** dell'**ipotesi statistica**

$$\boldsymbol{H_0}: \text{“}\textbf{la moneta è equilibrata}\text{”}$$

dipendono da valutazioni di tipo probabilistico in qualche modo vicine alla tecnica della dimostrazione per assurdo, ma anche affini al metodo che porta a **condanna** o **assoluzione** dell'imputato in un processo penale.

Le fasi della verifica di una ipotesi statistica

1) **H_0 : "la moneta è equilibrata"**

2) Definizione di un ***test*** che verrà effettuato, e delle modalità con cui un valore numerico "*Z*" verrà dedotto dal risultato del test.

3) Determinazione, con tecniche matematiche, di un intervallo " ***di accettazione***" in cui, **se è vera H_0** è molto probabile che si trovi *Z* (per esempio, con probabilità almeno 95%).

4) **Esecuzione del test e calcolo di** *Z*.

5) **Il verdetto.** ***Se Z non appartiene all'intervallo di accettazione***, l'esperimento ha prodotto un risultato che *a priori* era da ritenere *molto improbabile*. Questo fatto può essere spiegato in due modi opposti:

- $\boldsymbol{H_0}$ **è vera**, ma ugualmente si è potuto verificare qualcosa che, pur se poco probabile, non era impossibile.
- •• H_0 **è falsa**; le cose stanno diversamente, in un modo tale da rendere quel risultato più probabile di quanto valutato supponendo vera H_0 .

La regola di giudizio è di ritenere valida ***la seconda*** delle due possibili spiegazioni: se Z non appartiene all'intervallo di accettazione, H_0 viene ***rifiutata***, ossia giudicata ***falsa***.

Se Z appartiene all'intervallo di accettazione, l'esito dell'esperimento semplicemente **non appare in contrasto** con l'assunzione che H_0 sia vera; non vengono portati concreti elementi a supporto della verità di H_0, ma nemmeno contro di essa; la regola di giudizio è che in questo caso H_0 sia accettata o meglio, "non rifiutata".

Il criterio di condanna (rifiuto di H_0) **consiste nel considerare IMPOSSIBILE il verificarsi di qualcosa che, sotto l'ipotesi di innocenza** (verità di H_0)**, ha una probabilità molto bassa, non superiore a un livello prestabilito** α **detto SIGNIFICATIVITÀ. Spesso questo livello è fissato a 0,05 vale a dire 5%**

Il processo alla moneta.

H_0 : **Supponiamo** che la moneta sia equilibrata

Il test: lanciamo per 200 volte la moneta; indichiamo con ***t*** il numero di uscite di "testa".

Significatività e intervallo di accettazione:
$\alpha = 0{,}05$ (5%). Si calcola poi (anche a mano, con parecchia pazienza) che è 0,96 la probabilità che, **se vale** H_0 , si abbia $86 \le t \le 114$.

Esito del test: $t = 37$ non compreso nell'intervallo di accettazione (e assai lontano da questo!).
Il verdetto: H_0 **viene rifiutata.**

Il criterio di decisione che assumiamo consiste nel trattare

$$86 \leq numero\ teste \leq 114$$

come un **evento certo**, nel caso che H_0 sia vera.

Decidiamo cioè di **non credere** che un risultato del test fuori da questo intervallo possa avere luogo, se la moneta è equilibrata.

ERRORE DI PRIMA SPECIE

Consiste nel "***condannare un innocente***", ossia **rifiutare** l'ipotesi che la moneta sia equilibrata in un caso in cui la moneta lo è.

La probabilità di commettere l'errore di prima specie (nel caso in cui H_0 sia vera) è la ***significatività*** del test, cioè il 5%, nel nostro esempio.

Meno facile la stima della probabilità dell'*errore di seconda specie*, ossia "***assolvere un colpevole***".

4367 caramelle

382 rosse

2093 mangiate

Ipotesi nulla ("innocenza"): Il Re non ha specifiche preferenze riguardo al colore delle caramelle.

Ipotesi alternativa: il Re non gradisce le caramelle rosse. Di ciò abbiamo ***a priori*** qualche sospetto.

Test unilaterale: stabilito il valore α della significatività, determiniamo la regione di accettazione nella forma: "almeno *N* rosse".

Il test: facciamo scegliere al re 2093 caramelle; Z indica il numero di caramelle rosse.

Significatività e intervallo di accettazione:

Assumiamo un'altissima significatività:

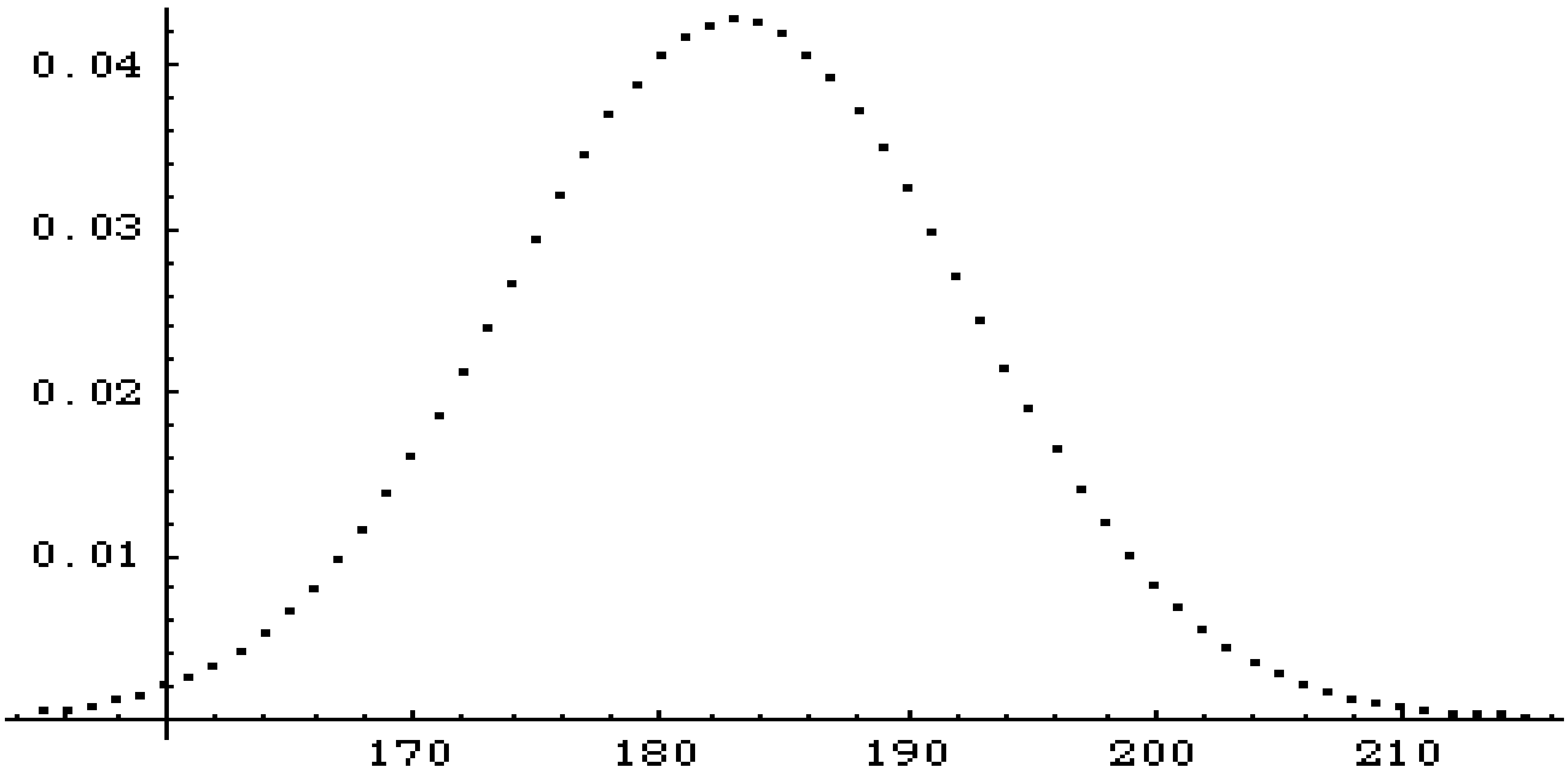
$$\alpha = 10^{-7} = 0{,}0000001 = 0{,}00001\%$$

Si calcola che il corrispondente numero N è **135**, ossia con probabilità **maggiore** di $1-10^{-7}=0{,}9999999$ almeno 135 caramelle estratte saranno rosse. L'intervallo di accettazione comprende quindi i valori $Z \geq 135$.

<u>Esito del test</u>: $Z = 0$ <u>non</u> compreso nell'intervallo di accettazione (e lontanissimo da questo!).
<u>Il verdetto</u>: H_0 **viene rifiutata.**

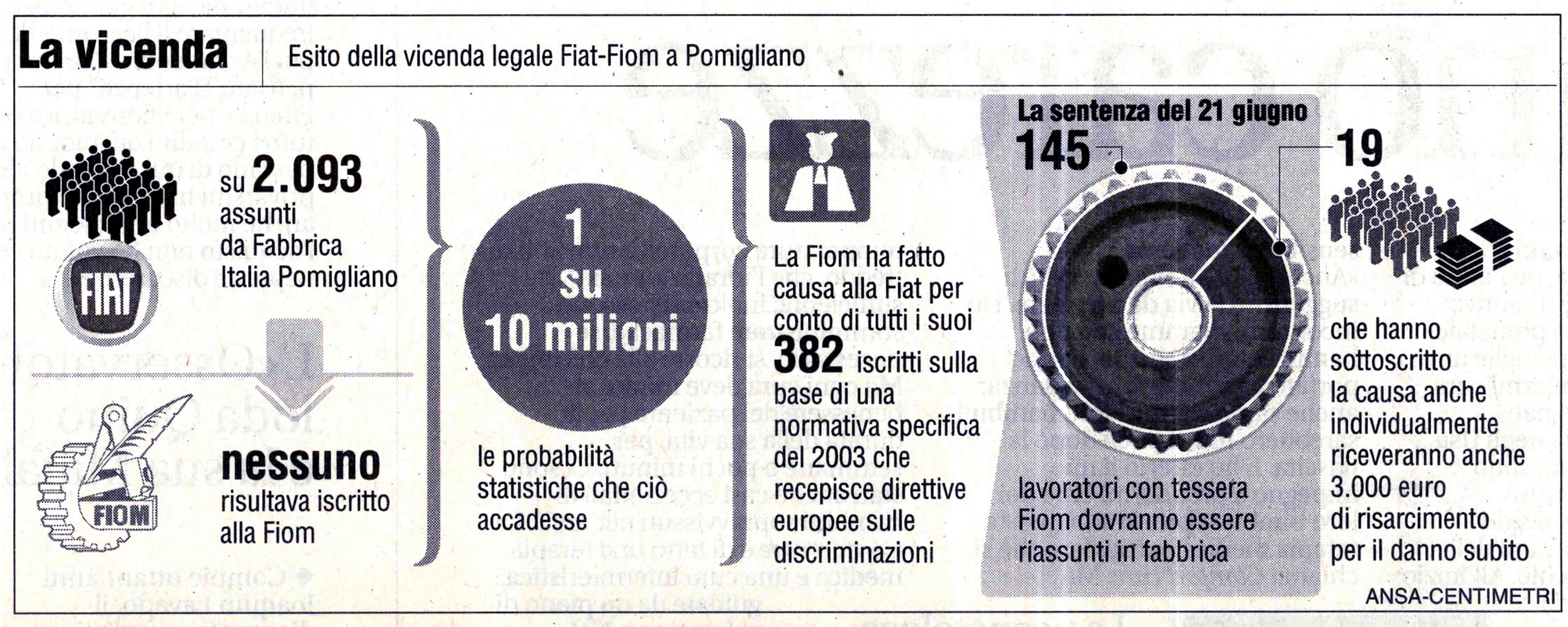
Conclusione: RIFIUTIAMO l'ipotesi di partenza, ossia che il re scelga le caramelle che intende mangiare indipendentemente dal loro colore.

L'altissima significatività assunta per il test e il risultato ottenuto danno la "quasi certezza" della bontà del verdetto. Il Re è stato giudicato secondo un principio estremamente garantista.



In ascissa: alcuni possibili valori per il numero di caramelle rosse presenti tra le 2093 estratte;

In ordinata: le corrispondenti probabilità.



(tra **4367** operai licenziati, **382** dei quali iscritti a *Fiom*, ne sono stati riassunti **2093**, in base a impegni assunti dall’azienda)

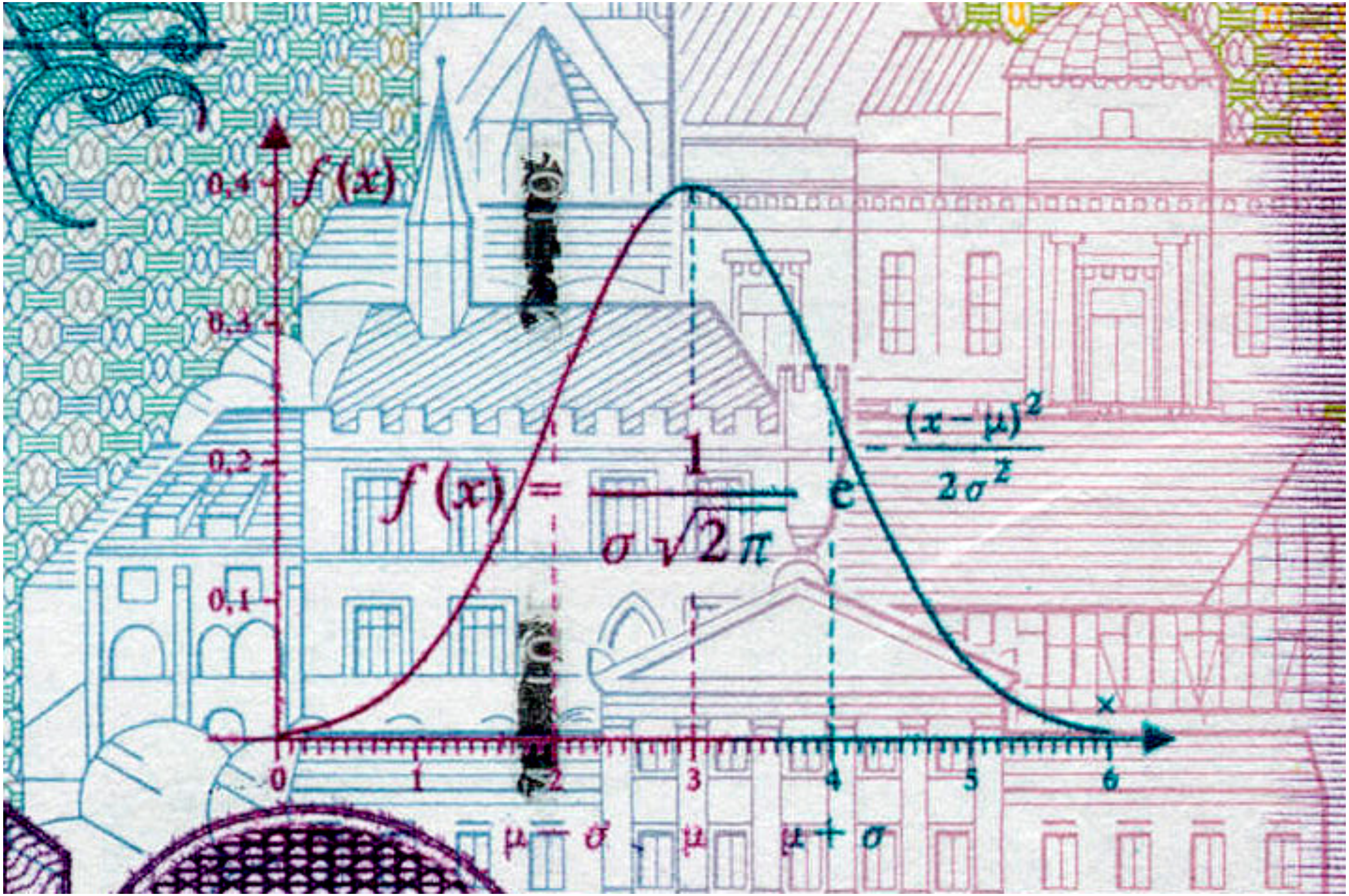
La tecnica: come si fa per determinare la regione di accettazione dell'ipotesi nulla.

Molti problemi del tipo di quelli esposti (ma non tutti), in cui si cerca un intervallo di valori nel quale con probabilità fissata a priori si troverà un risultato dedotto da un esperimento, sono riconducibili alla ***distribuzione normale standard***, oggetto degli studi del Principe dei Matematici:

Carl Friedrich Gauss

Ebbene, si dimostra che in un'ampia varietà di situazioni, valori (opportunamente trattati) dedotti da una serie "abbastanza numerosa" di esperimenti hanno una distribuzione di probabilità che segue la ***curva normale standard*** nota anche come ***curva di Gauss***, grafico della funzione

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$


$f(x)$
0,4
0,3
0,2
0,1
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
x
0
1
2
3
4
5
6
$\mu - \sigma$
μ
$\mu + \sigma$


DEUTSCHE BUNDESBANK
Banknote
10
f(x)
ZEHN DEUTSCHE MARK
10
1777–1855 Carl Friedr. Gauß
GU5672972S2

Un numero aleatorio Z si dice che ***è distribuito secondo la normale standard***, ovvero che "segue la distribuzione $N(0,1)$" se, fissati due numeri a, b con $a < b$, la probabilità che Z assuma un valore compreso tra a e b è uguale all'area delimitata dalla curva di Gauss tra gli estremi a e b.

Esistono tabelle che riportano i valori di queste aree; attualmente diversi programmi applicativi hanno la possibilità di fornire direttamente tali valori.

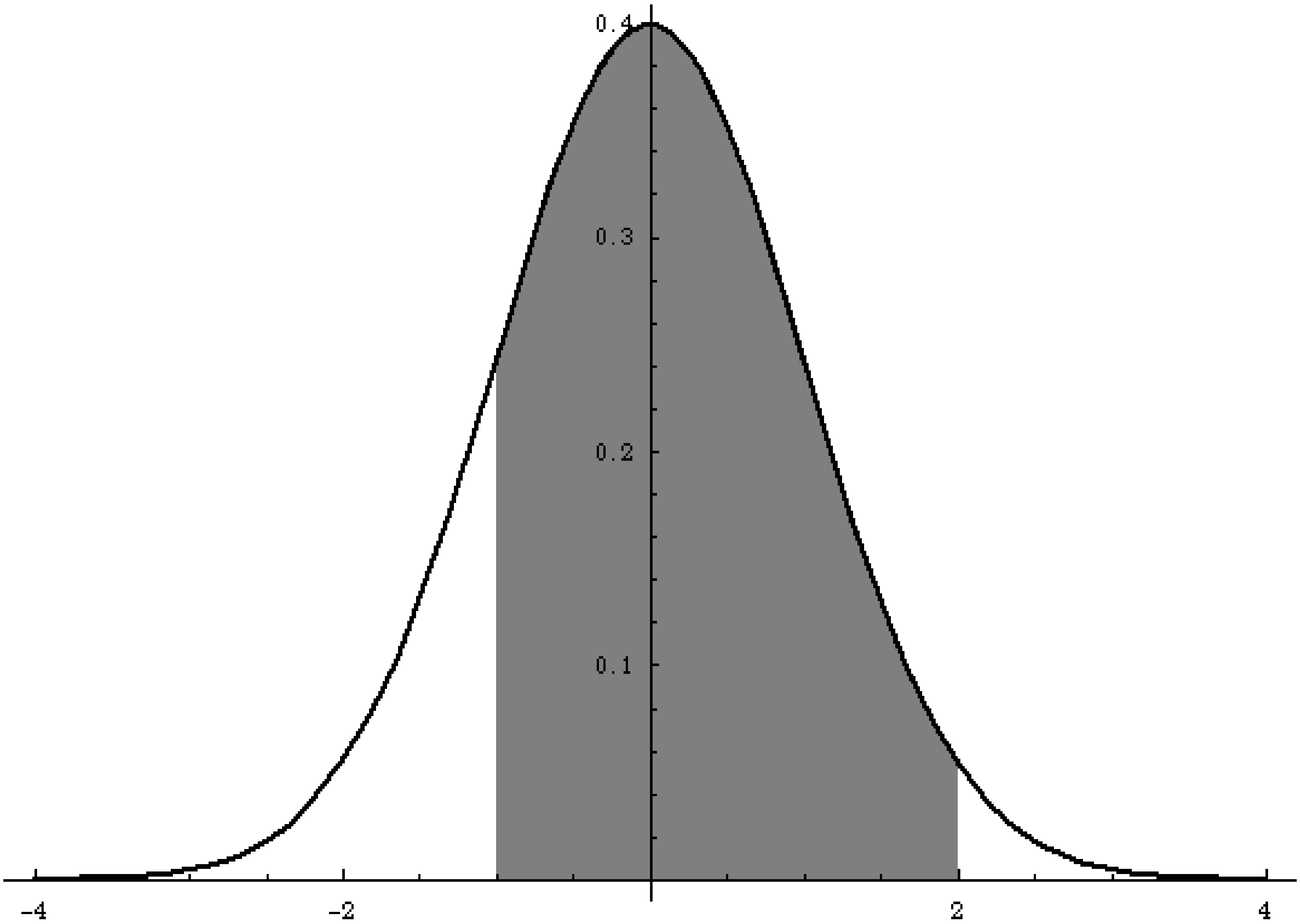
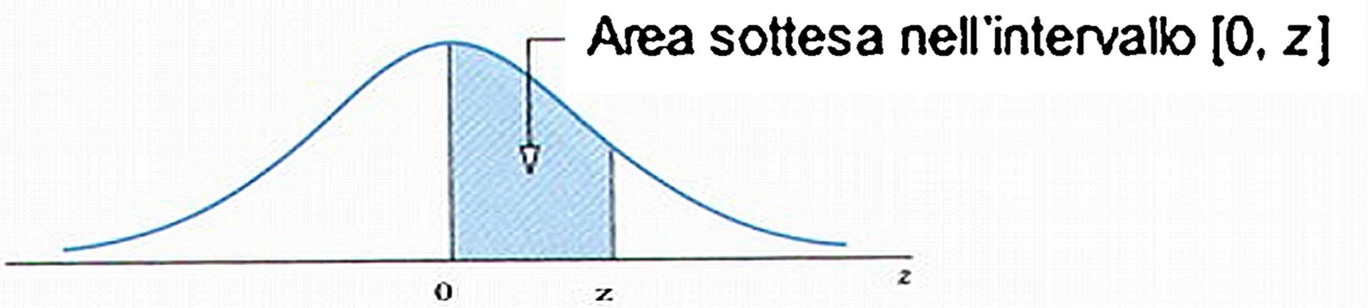
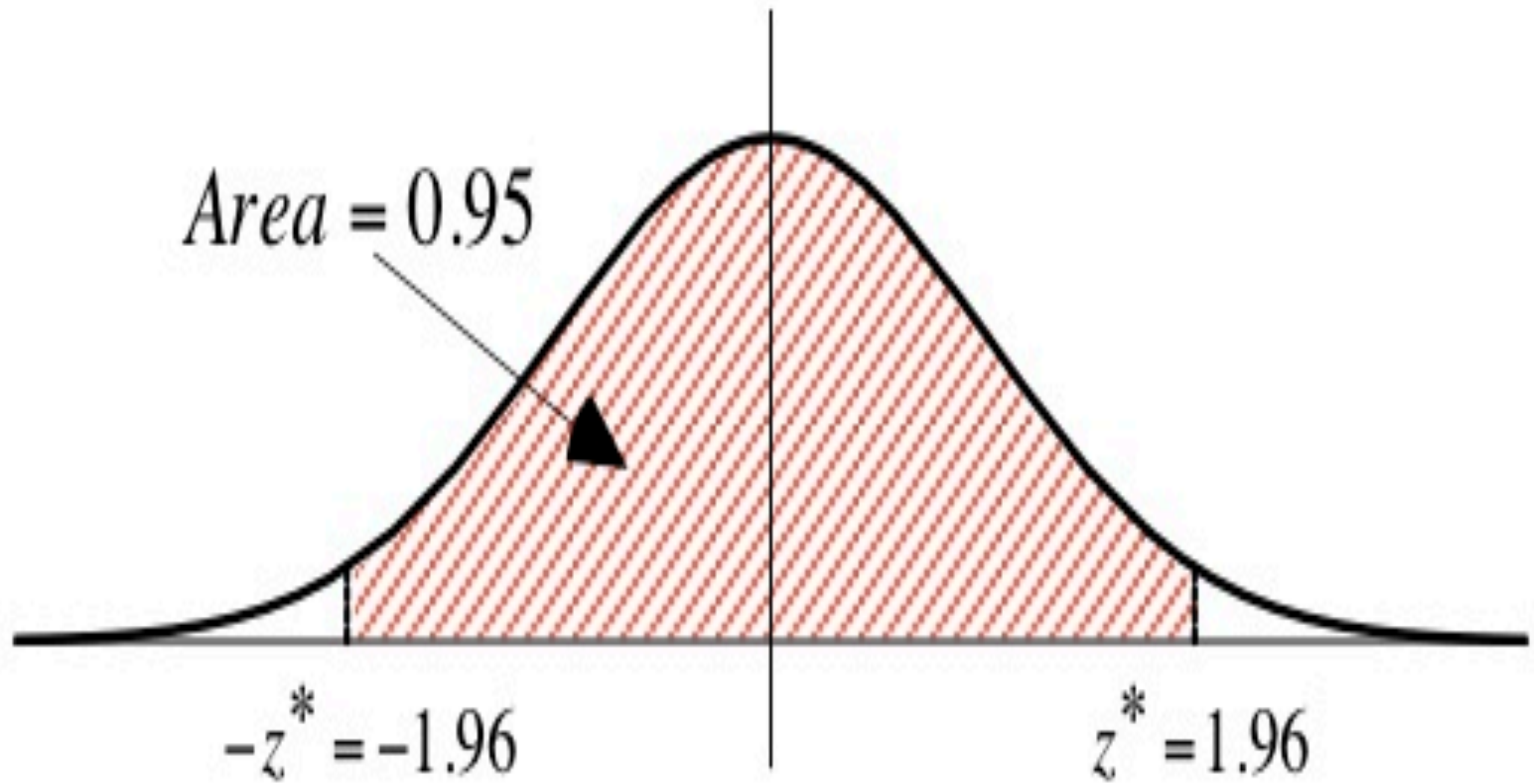

0.4
0.3
0.2
0.1
-4
-2
2
4

Tabella Distribuzione Normale Standard



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817



Area = 0.95
−z* = −1.96
z* = 1.96

-Per una percentuale,
ovvero per una probabilità:

se in una data popolazione un certo carattere è presente con una proporzione p_0 *e* ***f*** *è la* ***frequenza relativa*** *di quel carattere in un campione di* ***n*** *individui* $\left(f = \frac{\text{numero occorrenze del carattere osservato}}{n}\right)$

$(n \geq 30,\ n \cdot p_0 \geq 10)$, *allora il numero*

$$z = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

è distribuito come $N(0,1)$.

ovvero:

se in un esperimento è p_0 *la probabilità di ottenere un determinato risultato, chiamato "**successo**", e **f** è la frequenza relativa di successi che si ottengono ripetendo per **n** volte l'esperimento, cioè*

$$f = \frac{\text{numero di successi conseguiti}}{n}$$

$(n \geq 30,\ n \cdot p_0 \geq 10)$*, allora il numero*

$$z = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

è distribuito come $N(0,1)$.

Per esempio: $n = 200$ (lanci di una moneta), $p_0 = \frac{1}{2}$ (probabilità ***presunta*** dell'uscita di "testa"); $\boldsymbol{t}$ sia il numero aleatorio delle uscite di "testa" in 200 lanci; allora $f = t/n = t/200$. Il valore z "normalizzato" corrispondente al valore sperimentale $\boldsymbol{t}$ è

$$z = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{200}}} = \frac{\frac{t}{200} - \frac{1}{2}}{\sqrt{\frac{1}{800}}} = \frac{\sqrt{2}}{10}(t - 100)$$

Facilmente si calcola che

$$z = -1{,}96 \Leftrightarrow t = 86{,}14\,; \quad z = 1{,}96 \Leftrightarrow t = 113{,}86$$

-Per confrontare due percentuali, ovvero due probabilità:

Esempio tipico: verifica sperimentale dell'efficacia di un farmaco.

Un campione di n_1 persone affette da una malattia viene trattato con il farmaco da sperimentare; un altro campione di n_2 malati viene trattato con *placebo*. Nel primo gruppo si hanno g_1 guarigioni, nel secondo gruppo g_2 guarigioni; le frequenze relative di guarigioni in ciascuno dei due gruppi sono:

$$f_1 = \frac{g_1}{n_1}\,; \qquad f_2 = \frac{g_2}{n_2}$$

La ***ipotesi nulla*** (che questa volta si spera di dover rifiutare) è:

H_0: **"la probabilità di guarigione per chi ha ricevuto il farmaco è la stessa di chi ha ricevuto il placebo".**

in opposizione alla ***ipotesi alternativa*** (unilaterale)

H_1: **"la probabilità di guarigione per chi ha ricevuto il farmaco è *maggiore* di quella di chi ha ricevuto il placebo".**

Una **sensibile differenza** tra f_1 e f_2 (a favore della prima) avvalora l'ipotesi alternativa; al contrario, la vicinanza di queste due frequenze è compatibile con l'ipotesi nulla.

Si dimostra che, per campioni sufficientemente numerosi, ***il numero***

$$z = \frac{f_1 - f_2}{\sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}}$$

è distribuito come $N(0,1)$.

Per esempio, immaginiamo la seguente situazione:

	trattati	guariti
con farmaco	$n_1 = 80$	$g_1 = 60$
con placebo	$n_2 = 130$	$g_2 = 82$

Le frequenze relative delle guarigioni sono

$$f_1 = \frac{g_1}{n_1} = \frac{60}{80} = 0,75\ ; \quad f_2 = \frac{g_2}{n_2} = \frac{82}{130} = 0,63$$; quindi

$$z = \frac{f_1 - f_2}{\sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}} = \frac{0,75 - 0,63}{\sqrt{\frac{0,75 \cdot 0,25}{80} + \frac{0,63 \cdot 0,37}{130}}} = 1,85$$

Se la ***significatività*** viene fissata al 5%, l'intervallo di accettazione dell'ipotesi nulla (inefficacia del farmaco) è costituito dai valori di z compresi tra $-1,96$ e $1,96$.

I dati dell'esperimento hanno generato il valore $z = 1,85$ ***appartenente*** all'intervallo di accettazione.

Al livello di significatività del 5% ***non rifiutiamo*** l'ipotesi nulla; ciò significa che il modesto scarto tra le percentuali di guariti nei due gruppi ***viene attribuito al caso***, piuttosto che all'efficacia della medicina.

William Sealy Gosset (1876-1937) fu uno degli ideatori di queste tecniche per applicare la statistica alla Teoria delle decisioni. Matematico e chimico, alla dipendenze della fabbrica di birra ***Guinness*** di Dublino, egli affrontò il problema **se fosse utile oppure no essiccare preventivamente i semi dell'orzo prima della semina, per migliorare la produttività**. Pubblicò i suoi risultati sotto lo pseudonimo di ***Student***, perché il suo datore di lavoro temeva la divulgazione di segreti industriali; con questo pseudonimo egli è ricordato assai più che con il suo vero nome.