

# **Valutazione dei dati genetici**

- Il metodo statistico del chi-quadro

- Le previsioni statistiche sono tanto più valide quanto più grande è la dimensione del campione.
- Più grande è la dimensione del campione più piccola sarà la deviazione dai risultati attesi per effetto del caso.

*Esempio: Il rapporto sessi è 50%. Nelle singole famiglie il rapporto sessi di 1:1 è raro e devia molto dall'atteso, ma se prendo in considerazione l'intera popolazione il rapporto sessi (femmine/maschi) si avvicina a quello atteso.*

# Analisi del Chi-Quadro

- Il chi-quadro serve a stabilire se le differenze fra i dati osservati e quelli attesi possono essere attribuite al caso. Se le deviazioni dal rapporto atteso sono tali da non poter essere attribuite al caso abbiamo forse elaborato una ipotesi errata.
- **Passo 1:** Stabilire l' Ipotesi di partenza detta ipotesi 0/Nulla ( $H_0$ )

# Analisi del Chi-Quadro

- Quale sarebbe l' Ipotesi 0 in un incrocio tra monoibridi?
  - **Monoibridi**
    - I dati osservati **non sono** diversi dal rapporto atteso (3:1)
  - **Diibridi**
    - I dati osservati **non sono** diversi dal rapporto atteso (9:3:3:1)

# Chi- Quadro

- **Passo 2:** L' Equazione
- Noi usiamo l' equazione del chi-quadro per determinare se possiamo accettare l' ipotesi nulla (ossia se le differenze che osserviamo possono o meno essere attribuite al caso).

$$\chi^2 = \sum \frac{(\text{osservati} - \text{attesi})^2}{\text{attesi}}$$

- **Passo 3:** Calcolare il valore di chi-quadro per l'incrocio in esame
- Per ciascuna classe fenotipica, calcoliamo il valore di chi-quadro, poi sommiamo i valori di chi-quadro per ottenere un valore di chi-quadro che includa tutte le classi analizzate
- $\chi^2_{\text{totale}}$

# Passo 3: il Calcolo

<i>(a) Monohybrid Cross</i>					
<i>Expected Ratio</i>	<i>Observed (o)</i>	<i>Expected (e)</i>	<i>Deviation (o - e)</i>	<i>Deviation<sup>2</sup> (d<sup>2</sup>)</i>	<i>d<sup>2</sup>/e</i>
3/4	740	3/4 (1000) = 750	740 - 750 = -10	(-10) <sup>2</sup> = 100	100/750 = 0.13
1/4	260	1/4 (1000) = 250	260 - 250 = +10	(+10) <sup>2</sup> = 100	100/250 = 0.40
Total = 1000					$\chi^2 = 0.53$
					$p = 0.48$

• **Progenie osservata F2**

**740 piante con il carattere dominante**

**260 piante con il carattere recessivo**

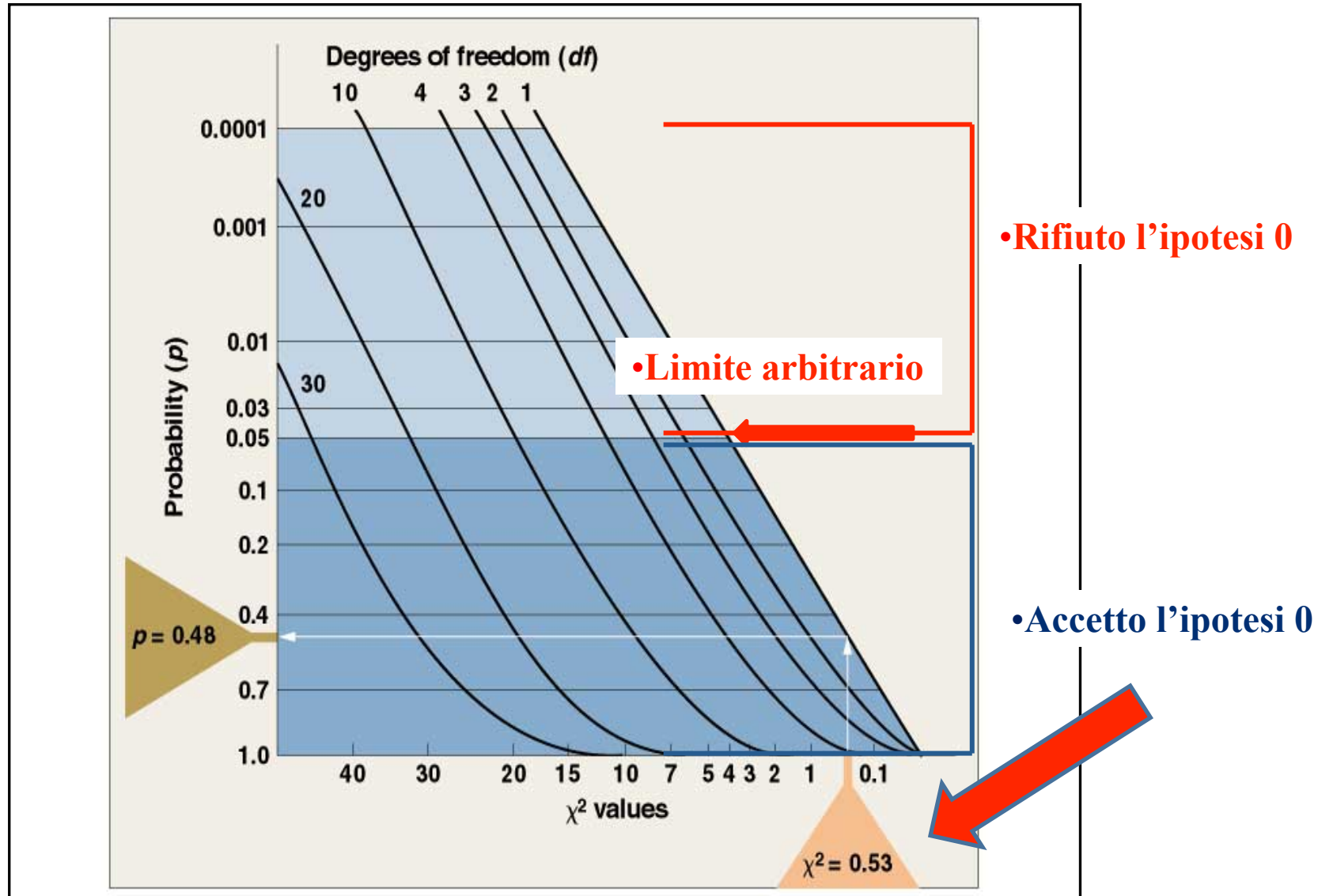
**Totale della progenie 1000**

**Il valore del Chi-quadro è 0.53**

- **Passo 4:** Determinare i gradi di libertà per quell' incrocio =  $n - 1$  dove  $n$  = numero di differenti classi (fenotipi in questo caso)
  - Per un incrocio monoibrido, fenotipi = 2
  - Gradi di libertà =  $2 - 1 = \mathbf{1}$

- **Passo 5** : Successivamente, dobbiamo determinare il valore- $p$  ossia il grado di confidenza con cui accetto o rifiuto l'ipotesi nulla.
  - In Genetica, quasi sempre si usa come valore limite :  $p = .05$  or 5% (grado di confidenza). Ossia il limite per accettare l'ipotesi 0 è che la deviazione dei nostri dati dall'atteso si osservi il 5% delle volte o più (effetto del caso).
  - Tale valore è arbitrario e potremmo usare valori più restrittivi  $p=.01$  o  $.001$
- Usiamo la tabella del Chi-quadro per trovare il valore di  $\chi^2$  tabulare

Chi-Square :questa tabella mostra i valori di probabilità che la deviazione dei dati osservati dagli attesi sia dovuta al caso. Se  $p$  è minore di 0.05 rifiuto l'ipotesi 0, se  $p$  è uguale o maggiore di 0.05 accetto l'ipotesi formulata.



# Chi-Square Probabilities

• Accetto l'ipotesi 0

• Rigelto l'ipotesi 0

		Probability ( $p$ )					
		0.90	0.50	0.20	0.05	0.01	0.001
<i>df</i>	1	0.02	0.46	1.64	3.84	6.64	10.83
	2	0.21	1.39	3.22	5.99	9.21	13.82
	3	0.58	2.37	4.64	7.82	11.35	16.27
	4	1.06	3.36	5.99	9.49	13.28	18.47
	5	1.61	4.35	7.29	11.07	15.09	20.52
	6	2.20	5.35	8.56	12.59	16.81	22.46
	7	2.83	6.35	9.80	14.07	18.48	24.32
	8	3.49	7.34	11.03	15.51	20.09	26.13
	9	4.17	8.34	12.24	16.92	21.67	27.88
	10	4.87	9.34	13.44	18.31	23.21	29.59
	15	8.55	14.34	19.31	25.00	30.58	37.30
	25	16.47	24.34	30.68	37.65	44.31	52.62
	50	37.69	49.34	58.16	67.51	76.15	86.60

$\chi^2$  values

**Tabella 13.1** Test del chi-quadrato utilizzato con i risultati del reincrocio per valutare l'ipotesi che i due geni sono indipendenti

(1) Fenotipi	(2) Numero osservato ( <i>o</i> )	(3) Numero atteso ( <i>e</i> )	(4) <i>d</i> (= <i>o</i> - <i>e</i> )	(5) <i>d</i> <sup>2</sup>	(6) <i>d</i> <sup>2</sup> / <i>e</i>
Parentali: (nero, normale e grigio, vestigiale)	2712	1618	1094	1.196.836	739,7
Ricombinante (nero, vestigiale e grigio, normale)	524	1618	-1094	1.196.836	739,7
Totale	3236	3236	3236		1479,4
(7) $\chi^2 = 1479,4$ (8) df 1					

**Tabella 11.5** Probabilità del chi-quadrato

df	Probabilità									
	0,95	0,90	0,70	0,50	0,30	0,20	0,10	0,05	0,01	0,001
1	0,004	0,016	0,15	0,46	1,07	1,64	2,71	3,84	6,64	10,83
2	0,10	0,21	0,71	1,39	2,41	3,22	4,61	5,99	9,21	13,82
3	0,35	0,58	1,42	2,37	3,67	4,64	6,25	7,82	11,35	16,27
4	0,71	1,06	2,20	3,36	4,88	5,99	7,78	9,49	13,28	18,47
5	1,15	1,61	3,00	4,35	6,06	7,29	9,24	11,07	15,09	20,52
6	1,64	2,20	3,83	5,35	7,23	8,56	10,65	12,59	16,81	22,46
7	2,17	2,83	4,67	6,35	8,38	9,80	12,02	14,07	18,48	24,32
8	2,73	3,49	5,53	7,34	9,52	11,03	13,36	15,51	20,09	26,13
9	3,33	4,17	6,39	8,34	10,66	12,24	14,68	16,92	21,67	27,88
10	3,94	4,87	7,27	9,34	11,78	13,44	15,99	18,31	23,21	29,59
11	4,58	5,58	8,15	10,34	12,90	14,63	17,28	19,68	24,73	31,26
12	5,23	6,30	9,03	11,34	14,01	15,81	18,55	21,03	26,22	32,91
13	5,89	7,04	9,93	12,34	15,12	16,99	19,81	22,36	27,69	34,53
14	6,57	7,79	10,82	13,34	16,22	18,15	21,06	23,69	29,14	36,12
15	7,26	8,55	11,72	14,34	17,32	19,31	22,31	25,00	30,58	37,70
20	10,85	12,44	16,27	19,34	22,78	25,04	28,41	31,41	37,57	45,32
25	14,61	16,47	20,87	24,34	28,17	30,68	34,38	37,65	44,31	52,62
30	18,49	20,60	25,51	29,34	33,53	36,25	40,26	43,77	50,89	59,70
50	34,76	37,69	44,31	49,34	54,72	58,16	63,17	67,51	76,15	86,66

← Accettare | Rifiutare →  
al livello di 0,05

Fonte: Estratto dalla Tabella IV in *Statistical Tables for Biological, Agricultural, and Medical Research* di Fisher e Yates, 6<sup>a</sup> ed., 1974. Ristampato per gent. conc. di Addison Wesley Longman Ltd,

rattere era un dominante semplice. In questa sezione ver

# Interpretazione dei dati

- Così, a un valore- $p$  di 0.05 and 1 grado di libertà
  - Il valore limite ad 1 grado di libertà è  $\chi^2$  tabulare = 3.84
  - Il valore di  $\chi^2$  calcolato = 0.53
- Il valore calcolato mi dà  $p >$  di 0.05
  - Quindi, accettiamo  $H_0$
  - In altre parole la deviazione dei dati osservata è attribuibile al caso con un grado di confidenza superiore al 5%.

## **Esercizio 1**

**Una pianta ha fiori porpora o bianchi ed i baccelli lisci o irregolari.**

**Determinare l'allele dominante per i due caratteri e indicare il genotipo dei genitori per il seguente incrocio.**

**Porpora irregolare X porpora irregolare**

**94 porpora irregolare**

**32 bianco irregolare**

**28 porpora liscio**

**11 bianco liscio**

**Applicate il test del CHI quadrato per verificare la vostra ipotesi.**

# Interpretazione dei dati

- $H_0$  = Il rapporto misurato **non è significativamente diverso** dal rapporto atteso
  - Le deviazioni osservate sono quindi attribuibili al caso, l'ipotesi  $H_0$  può essere accettata.

## **Esercizio 2**

**I moscerini selvatici hanno ali corte e colore del corpo marrone.**

**Un incrocio fra un moscerino omozigote per l'allele ali vestigiali (lunghe) ed un moscerino con corpo ebony (nero). Produce una F1 selvatica. Incrociando la F1 con moscerini con ali vestigiali e corpo nero si ottenne nella F2 la seguente progenie:**

**565 selvatici**

**498 con ali vestigiali**

**546 con corpo ebony**

**527 con ali vestigiali e corpo ebony**

**Impostate i genotipi**

**Fate la vostra ipotesi**

**Applicate il test del CHI quadrato per verificare la vostra ipotesi.**

Come vedremo in seguito alcune malattie umane sono legate o influenzate dal sesso.

Pertanto nella consulenza genetica si rende necessario il calcolo del rapporto sessi e della associazione con la malattia.

Es: l'emofilia

- Calcolare i valori di probabilità del rapporto sessi nella progenie di una famiglia composta da 5 figli.
- $P_{\text{maschio}} = P_{\text{femmina}} = 1/2$
- $P_{\text{maschio}} = p \quad P_{\text{femmina}} = q \quad n = 5$
- $(p + q)^n$
- $p^5 + p^4q + p^3q^2 + p^2q^3 + pq^4 + q^5$
- Le possibili combinazioni sono date dall'espansione del binomio  $(p + q)^5$

- $p^5 + p^4q + p^3q^2 + p^2q^3 + pq^4 + q^5$

- I coefficienti si ottengono mediante il triangolo di Pascal.

- $n=1$   $1$

- $n=2$   $1$   $2$   $1$

- $n=3$   $1$   $3$   $3$   $1$

- $n=4$   $1$   $4$   $6$   $4$   $1$

- $n=5$   $1$   $5$   $10$   $10$   $5$   $1$

- Quindi:  $p^5 + 5 p^4q + 10 p^3q^2 + 10 p^2q^3 + 5 pq^4 + q^5$

- $p^5 + 5 p^4 q + 10 p^3 q^2 + 10 p^2 q^3 + 5 p q^4 + q^5$
- La probabilità che in una famiglia di 5 figli nascano 5 maschi è  $(1/2)^5$
- Qual è la probabilità che in una famiglia di 5 figli solo 1 dei figli sia di sesso diverso?
-

# Esercizi

- In una famiglia di 3 figli, calcolare la probabilità di avere due femmine ed un maschio.
- Tre figli dello stesso sesso.

Con quale probabilità una famiglia con 8 figli sarà composta da 6 maschi e 2 femmine?

In alternativa al metodo appena studiato possiamo applicare questa formula:

$$\frac{N!}{S! \times T!} p^s q^t$$

**N = totale eventi = 8**

**S = totale maschi**

**T = totale femmine**

**p = probabilità evento a**

**q = probabilità evento b**

- Si considerino tre semi di pisello gialli lisci, indicati con A, B, C. Ogni pisello germina e dà luogo ad una pianta. Le tre piante sono poi incrociate con una pianta nata da un seme verde rugoso. Da ciascun incrocio si ottengono esattamente 100 piselli, suddivisi nelle seguenti classi fenotipiche:
  - A 51 gialli lisci  
49 verdi lisci
  - B 100 gialli lisci
  - C 24 gialli lisci  
26 gialli rugosi  
25 verdi lisci  
25 verdi rugosiQuale è il genotipo dei semi A B C ?

## **Esercizio**

**(Applicate la formula con i fattoriali)**

**La galattosemia è una malattia dovuta ad un allele recessivo. Susan ed il marito sono entrambi eterozigoti per il gene della galattosemia.**

**Susan è incinta di due gemelli dizigotici.**

**Che probabilità hanno di essere due sorelle galattosemiche?**

**Se i due gemelli fossero identici che probabilità avrebbero di essere due sorelle galattosemiche?**

**Se la coppia avesse tre figli qual è la probabilità che due siano colpiti da galattosemia ed uno no?**

**Se la coppia avesse tre figli qual è la probabilità che il primo sia sano ed i successivi siano colpiti da galattosemia?**

- Si suppone che una certa pianta sia diibrido per due geni che assortiscono indipendentemente. H/h R/r
- (H foglie pelose h foglie lisce)
- (R ovario sferico r ovario allungato)
- I numeri di individui per classe fenotipica indicati di seguito sono coerenti con i risultati attesi da un'autofecondazione?
- Foglie pelose, ovario sferico 178
- Foglie pelose, ovario allungato 62
- Foglie lisce, ovario sferico 56
- Foglie lisce, ovario allungato 24

- Un genetista ha tre linee pure di mais con i seguenti genotipi: aaBBCC      AAbbCC  
AABBcc. Tutti i fenotipi recessivi fanno aumentare il valore del mais sul mercato per cui il genetista vuole ricavare una linea pura aabbcc
- Definite un efficace programma di incroci. Qual è il modo migliore per ottenere il risultato desiderato?