

LEZIONE N° 8 GENETICA

13/04/2023

GENETICA DI TIPO QUANTITATIVO (HARDY-WEINBERG)

[PRINCIPIO DI HARDY E WEINBERG, Hardy (matematico britannico) e Weinberg (medico tedesco) ricercano le cause dell'origine e del mantenimento della variabilità genetica. Essi cercano di comprendere come sia gli alleli dominanti che i recessivi riuscissero a rimanere all'interno delle popolazioni. Loro volevano capire il motivo per cui i dominanti non eliminano i recessivi. Gli scienziati, infatti, dimostrarono che la ricombinazione genetica non modifica di per sé il pool genico stesso].

Ma prima parliamo della genetica di tipo **quantitativo** perché molte malattie sono piuttosto rare con determinate frequenze in specifiche popolazioni (come la trasmissione della PKU (Fenilchetonuria) nei cinesi, negli irlandesi [1/16.000] o negli italiani [1/2.500]).

L'**eterogeneità** è dovuta alla diversa distribuzione di frequenza allelica nelle popolazioni, alcune patologie si diffondono infatti in maniera differente rispetto ad altre e dunque per ogni consulenza genetica che si richiede, bisogna tener conto del background etnico dal quale deriva il nostro paziente.

Ormai sono descritte nei diversi database le diverse frequenze alleliche (delle patologie) e le indicazioni di trasmissione della patogenicità delle specifiche popolazioni (quindi si sta parlando specificatamente della genetica di popolazione).

LA GENETICA DI POPOLAZIONE

La genetica di popolazione studia l'ereditarietà di trasmissione di determinati geni.

Studia dunque la **distribuzione dei geni** nella popolazione e dei **fattori che mantengono o modificano la frequenza** dei geni di generazione in generazione.

Individui della stessa popolazione condividono coppie di alleli costanti che descrivono e compongono il pool genetico (la genetica studia dunque il genotipo ed il fenotipo di una popolazione rappresentativa della nostra statistica).

Nella genetica di popolazione contiamo gli alleli di una specifica popolazione, estrapolando e studiando uno specifico campione.

Studia inoltre la variazione delle frequenze alleliche nel corso del tempo. Distinguiamo dunque la frequenza allelica (numeri di quegli specifici alleli sul totale di alleli) e la frequenza genotipica (numeri di individui manifestanti quel gene sul totale di individui). RICORDA CHE: Gli organismi diploidi hanno due alleli di ogni gene (uno del padre e uno della madre).



Dunque, tali organismi partecipano alla trasmissione del pool genico. Il **pool genico** è la somma di tutti i geni che rappresentano quello specifico carattere nella popolazione campione. [può anche essere espresso in percentuale].

[All'inizio del XX secolo gli scienziati si chiedevano come mai nelle popolazioni naturali alcuni alleli aumentassero di frequenza da una generazione ad un'altra, mentre altri diminuivano]. Pertanto, si distinguono: # **Alleli favorevoli**, con maggiore probabilità di sopravvivere e di riprodursi, tali alleli hanno maggiori probabilità di essere

presenti in proporzione crescente nelle generazioni successive; # **Alleli sfavorevoli**, con minore probabilità di sopravvivere e di riprodursi, la loro presenza nelle generazioni successive risulterà ridotta o assente. Si parla poi di **Fitness Darwiniana** per indicare il successo riproduttivo, ovvero il numero di discendenti che sopravvivono, che determina quanti alleli del genotipo di un individuo possono essere presenti nelle generazioni successive.

FITNESS DARWINIANA

Misura della selezione naturale, ossia dell'efficienza riproduttiva di un **genotipo** in un dato ambiente ed in un preciso momento.

Viene calcolata come media dei tassi di sopravvivenza e riproduzione degli individui aventi lo stesso genotipo, coesistenti in uno stesso ambiente e in un dato momento.

La fitness realmente utile è quella **relativa**, ossia il parametro che mette in relazione un genotipo con gli altri possibili.



Ritornando al concetto di genetica quantitativa e in particolare del principio di Hardy-Weinberg.

SAPPI CHE: Ancor prima di affrontare il principio secondo cui si può ricavare il valore atteso di una popolazione, bisogna tener presente la visualizzazione diretta della popolazione per poi confrontare i due risultati. La conta diretta può essere effettuata per qualsiasi locus i cui alleli siano codominanti.

RICORDA CHE: Un esempio di codominanza sono i gruppi sanguigni umani. Nella codominanza distinguiamo tre differenti fenotipi dagli stessi genotipi. Il tutto si basa su una conta diretta di frequenze alleliche; per quanti sono gli individui, avremo come alleli totali il doppio degli individui (in quanto parliamo di individui diploidi). Avremo due tipi di conte dirette:

1. Conta genotipica: conta quanti genotipi vi sono sul totale;
2. Conta allelica o genica: conta quanti alleli vi sono sul totale.

Dunque, in condizioni semplici queste frequenze possono essere stimate mediante conteggio diretto. Ad esempio, in una popolazione campione analizzata, in cui ricaviamo il genotipo, da questo possiamo ricavare sia la frequenza genica/allelica, sia la frequenza genotipica. Quando non vi è codominanza, in caso di eterozigosi il dominante nasconde il recessivo, si utilizzano le formule probabilistiche di base per ricavare la frequenza allelica.

Immaginiamo che in un locus abbiamo due alleli, etichettati "A" e "a". Se nella popolazione, conoscendo la frequenza degli alleli A e a, potremo chiamarli rispettivamente "p" e "q". Da questa base possiamo determinare le frequenze di popolazione attese di ciascun genotipo AA, Aa e aa.

	A (p)	a (q)
A (p)	AA (p^2)	Aa (pq)
a (q)	Aa (pq)	aa (q^2)

Le frequenze dei genotipi nella progenie sono quindi:

genotipo:	AA	Aa	aa
frequenza:	p^2	$2pq$	q^2

La relazione tra le frequenze alleliche e le frequenze genotipiche è stata stabilita indipendentemente da Hardy e Weinberg, ed è chiamata principio di Hardy-Weinberg.

$p^2 + 2pq + q^2 = 1$		
AA	Aa	aa
p^2	$2pq$	q^2

Questo non è altro che un quadrato di binomio, riscrivibile come :

$$p + q = 1$$

[Ad esempio, supponiamo che la frequenza , p , dell'allele A nella nostra popolazione sia 0,7; ciò significa che il 70% degli spermatozoi nella popolazione deve avere l'allele A , così come il 70% degli ovuli].

Poiché la somma delle frequenze p e q deve essere 1, il 30% degli ovuli e degli spermatozoi deve portare l'allele a (cioè $q = 0,30$).

La probabilità che uno spermatozoo portatore di A si unisca con un ovulo portatore A , è data dal prodotto delle frequenze alleliche: $p \times p = p^2 = 0,49$

E la frequenza degli eterozigoti nella popolazione? $2pq = 2 \times 0,7 \times 0,3 = 0,42$

[Ad esempio, supponiamo che in una popolazione naturale che vi siano presenti 3 fenotipi facilmente distinguibili dovuti a 2 alleli del locus A :

AA (452 individui); Aa (43 individui); aa (2 individui).

Da cui possiamo ricavare le frequenze genotipiche:

$$f(AA) = 452/497 = 0.909 \text{ (pari a } p^2)$$

$$f(Aa) = 43/497 = 0.087 \text{ (pari a } 2pq)$$

$$f(aa) = 2/497 = 0.004 \text{ (pari a } q^2)$$

La somma è 1 e le singole percentuali descrivono quantitativamente il pool genico per quel locus all'interno del gruppo di individui che sto considerando.

La struttura genetica di una popolazione viene espressa in termini di frequenza allelica (o genica) invece che genotipica (tenendo conto che gli alleli sono meno dei genotipi).

È più vantaggioso descrivere il pool genico in termini di frequenza degli alleli calcolando il numero di copie di un allele sul totale di alleli nel pool genico].

Le condizioni per avere una popolazione adeguata all'equilibrio di **HW** devono essere:

1. Grande popolazione. Ciò è necessario per poter applicare la legge dei grandi numeri e quindi le frequenze sono praticamente coincidenti con le probabilità. Sorprendentemente una popolazione di poche centinaia di individui è sufficiente, anche se sono possibili fluttuazioni (ma improbabile).
2. Assenza di immigrazione ed emigrazione. In questo modo il patrimonio genetico è influenzato solo dalla sua dinamica interna.
3. Panmissia (incrocio casuale). Significa che la probabilità che due individui si incrocino non è influenzata dal fenotipo del carattere in questione. In questo modo è come se i geni di tutti gli individui sono stati mescolati nel pool genetico e selezionati casualmente per creare genotipi di nuovi individui. Manca la Panmissia, ad esempio, nel caso di forti preferenze matrimoniali all'interno di caste chiuse, soprattutto se di origine etnica.
4. Non selezione. Il successo riproduttivo medio degli individui (chiamato anche fitness) non deve essere influenzato dal genotipo del carattere in questione. I due (o più) alleli devono avere quindi la stessa probabilità una volta presenti, di essere trasmessi alle generazioni successive.
5. Non mutazione. Ovviamente le mutazioni alterano la composizione del patrimonio genetico delle nuove generazioni. Tuttavia, sono eventi rari.

L'equilibrio di **HW** afferma che in presenza di tali condizioni, le frequenze alleliche e geniche rimangono costanti nel tempo.

Tali cinque condizioni non saranno mai perfettamente rispettate nella popolazione naturale, ma la legge di **HW** è fondamentale come punto di partenza per gli studi di medicina, come nel caso dello studio di genetica di popolazione e per rintracciare eventuali errori di genotipizzazione all'inizio di uno studio di genetica medica.

Quando si fa uno studio di genetica medica e si parte dallo studio della popolazione, la prima cosa da chiedersi è se essa sia in equilibrio di Hardy-Weinberg; quindi, vi sono delle formule che dicono che si è in equilibrio se:

Hardy-Weinberg Equations

$$p^2 + 2pq + q^2 = 1$$

p = frequency of the dominant allele
in a population

$$p + q = 1$$

q = frequency of the recessive allele
in a population

Gli estremi sono caratteristiche molto differenti; la possibilità di un accoppiamento fatto per caratteristiche genotipiche è difficile, parliamo di caratteristiche estetiche, ma anche caratteriali etc. ma questo non corrisponde al fatto che questa persona ha quel determinato enzima.

È vero che la scelta non casuale disturba l'equilibrio, ma dal punto di vista genico non è particolarmente rilevante poiché un individuo non porta il proprio genotipo "scritto in fronte", quindi in realtà posso avere un'affinità morfologica, estetica, culturale, ma ciò non dice nulla sul genotipo.

La cosa che può essere poco casuale è la **consanguineità, l'incrocio**, cioè la possibilità nelle malattie autosomiche recessive, quanto più spesso si può manifestare il carattere "malattia" poiché l'allele deleterio ha più possibilità di esprimersi in doppia dose e quindi di rivelarsi.

Questo, ad esempio, è l'effetto dell'*Inbreeding*: [si ha una prima generazione con un certo quantitativo di eterozigoti e di omozigoti dominante e recessiva, ma ad ogni generazione diminuisce la quantità di eterozigoti, si formeranno più omozigoti dominanti e più omozigoti recessivi, che probabilmente verranno eliminati nella popolazione]. Quindi **nell'inbreeding**, in questo caso, **si scoprono gli alleli rari recessivi, deleteri o letali**, aumenta l'omozigosi della popolazione ma non cambiano le frequenze alleliche, cambia solo la frequenza del genotipo.

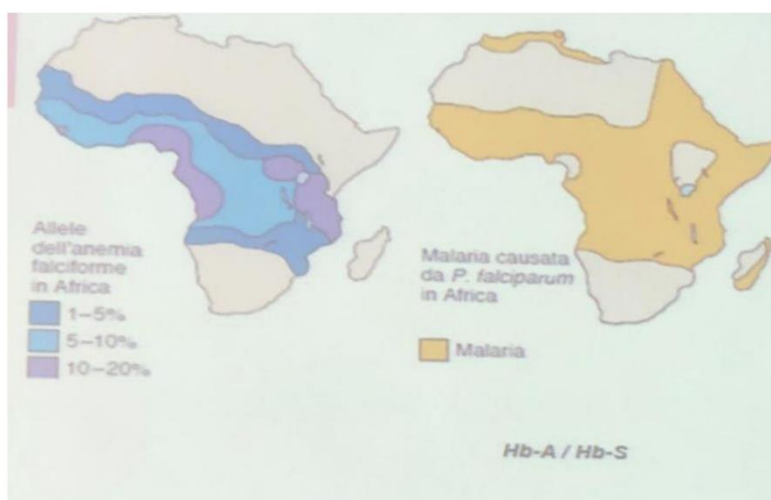
Quindi, in caso di matrimonio tra consanguinei, l'accoppiamento avviene in modo non casuale tra individui che presentano un simile genotipo, per cui maggiore la percentuale dei loro figli sarà omozigote per un allele o per un altro, maggiore sarà il cambiamento della frequenza genotipica e non della frequenza allelica, perciò la consanguineità, **da sola**, non modifica le sequenze alleliche ma cambierà la distribuzione genotipica.

La consanguineità da sola non modifica le frequenze alleliche ma, alterando l'unione dei geni a formare i genotipi, modifica la distribuzione genotipica

P:	Aa x Aa		
F1:	25% AA	50% Aa	25% aa
F2:	37.5% AA	25% Aa	37.5% aa
F3:	43.75% AA	12.5% Aa	43.75% aa

Meno eterozigoti e più omozigoti ad ogni generazione

Le mutazioni sono un evento molto raro, ma possibile. È l'unico mezzo attraverso il quale una specie possa acquisire nuovo materiale genetico de novo. È uno di quei 5 elementi che possono variare l'equilibrio in una popolazione naturale, l'insorgenza di una mutazione favorisce il cambiamento delle frequenze alleliche e genotipiche ma non tutte le mutazioni sono letali o dannose, alcune possono dare vantaggi. Le mutazioni dannose sono quelle che vengono eliminate a meno che non si abbia il cosiddetto vantaggio dell'eterozigote (portatore). La selezione naturale è un altro di quegli elementi che devia l'equilibrio della popolazione. È un processo della natura che seleziona la forma più adeguata a sopravvivere e a riprodursi in un determinato ambiente: se un allele fornisce un vantaggio ad un individuo che quindi porta la sua frequenza allelica ad aumentare si ha una selezione positiva, ma si può parlare anche di selezione negativa quando la frequenza degli alleli che portano a svantaggi per la popolazione, diminuisce ad ogni generazione. Questo è il parametro che si utilizza per determinare la selezione naturale insieme alla fitness del genotipo, che è la capacità di riprodursi e quindi di trasmettere l'allele (e il proprio patrimonio genetico) con successo, nelle generazioni successive (misura della capacità di un individuo di riprodursi).



Vantaggio dell'eterozigosi: un esempio ne è la talassemia [L'anemia mediterranea, anche nota come beta-talassemia major o anemia di Cooley, è una **malattia del sangue ereditaria** molto grave causata da un difetto genetico che **provoca la distruzione dei globuli rossi**] e le emoglobinopatie, che sono distribuite soprattutto in Africa ed hanno permesso agli individui eterozigoti la sopravvivenza alla malaria. Vi è infatti una sovrapposizione tra le zone in cui sono presenti casi di anemia falciforme e zone affette da malaria. Gli omozigoti

recessivi sono malati di anemia falciforme; negli omozigoti dominanti il ***Plasmodium Falciparum*** (protozoo della malaria) si può insediare e riprodurre; **gli eterozigoti, invece, non sono né malati di anemia falciforme (aut. Rec.) né possono contrarre la malattia della malaria.** → l'eterozigote quindi persiste nella popolazione. Inoltre, gli eterozigoti (per altri alleli) sono resistenti alla febbre tifoide e alla salmonella.

Nell'anemia falciforme abbiamo due emoglobine, [di tipo A e di tipo S, quest'ultima è il tratto falcemico ed è dovuto al cambiamento, nel locus dell'emoglobina, di un amminoacido, della valina al posto dell'acido glutammico; questo cambiamento cambia l'affinità dell'emoglobina per l'ossigeno (in quella mutata sarà ridotta)], **quindi gli individui omozigoti** per la mutazione **saranno affetti da anemia falciforme** e i loro globuli rossi assumeranno la tipica forma "a falce", in quanto, al loro interno, la pressione parziale dell'ossigeno è ridotta e può portare a emolisi; ciò non succede negli individui eterozigoti. Nelle zone di incidenza della malaria (anche mediterranee) si ha dunque la prevalenza dell'allele falcemico. Sebbene la selezione agisca contro l'omozigosi recessiva, il vantaggio dell'eterozigote mantiene l'allele nella popolazione. In zone geografiche dove non è presente questo vantaggio (non è presente la malaria) l'allele non sarà mantenuto. In genetica, più che di migrazione, si parla di flusso genico, la prima si ha quando gli individui si spostano da un'area geografica ad un'altra e si ha la determinazione di un flusso genico perché si ha la graduale diffusione di alleli da una popolazione ad un'altra.

Questo flusso può cambiare il pool genico della popolazione perché i migranti introducono nuovi alleli, può quindi cambiare le frequenze alleliche se le frequenze alleliche dei migranti e quelle della popolazione sono diverse, ma riduce le differenze genetiche tra le popolazioni. Quindi nelle generazioni successive le divergenze genetiche si standardizzeranno. Parliamo quindi anche di deriva genetica, che indica una fluttuazione, una variazione casuale delle frequenze alleliche e geniche passando ad una generazione successiva. In seguito ad una catastrofe, ad un brusco cambiamento, ad una diminuzione della popolazione si potrà presentare la fissazione di un allele e possiamo parlare di un insieme di variazioni dovute al caso che conduce ad una perdita di variabilità genetica all'interno di una popolazione. Esempi di deriva:

1. **Collo di bottiglia:** quando la popolazione si riduce così tanto che restano pochi individui a garantire la sopravvivenza della specie. In questo caso la distribuzione degli alleli e dei geni dipenderà dal genotipo di pochi e la variabilità sarà ridotta
2. **Effetto del fondatore:** quando la popolazione è fondata su un numero ristretto di individui (fondatori).
[Un esempio pratico è stato dato dagli eventi successivi alla scoperta dell'America].

Per le popolazioni di natura umana, l'**equilibrio di HW** è solo teorico, è improbabile che tutti i criteri siano soddisfatti, e ha dei limiti in quanto l'evoluzione di una popolazione è favorita proprio da variazioni dell'equilibrio genico (mutazioni, selezione naturale etc. , immigrazione...) ma viene utilizzato negli studi di medicina perché permette di predire fenotipi patologici in una popolazione, oltre a permettere di valutare l'effetto di situazioni reali sulle frequenze alleliche. Si può quindi fare una discriminazione tra frequenze attese e frequenze osservate. **Se le frequenze attese sono molto diverse da quelle osservate, può esserci un errore di calcolo** o ci si può trovare davanti ad una popolazione non in equilibrio quindi la differenza non è dovuta al caso. L'applicazione di **HW permette di calcolare frequenze genotipiche dalle frequenze alleliche** (e viceversa), permette di calcolare, nelle malattie autosomiche recessive, frequenze del portatore ed è molto importante (per esempio, nella fibrosi cistica, nel daltonismo etc.) **e permette di stabilire se una popolazione si trovi in equilibrio** (quindi se si può andare avanti nello studio di calcoli statistici). [Esempio: calcola frequenze genotipiche dalle frequenze alleliche]

n : allele recessivo \rightarrow frequenza 0,4. Trovare la frequenza genotipica

$q = 0,4$. $p = 1 - 0,4$ (secondo il binomio di HW) = 0,6

$NN = p^2 = 0,36$ $nN = 2pq = 2 \times 0,4 \times 0,6 = 0,48$. $Nn = q^2 = 0,16$

Esempio contrario con i gruppi sanguigni:

Gruppi sanguigni MN

MM	1787 individui
MN	3039 individui
NN	1303 individui
Totale	6129

$$\text{Frequenza allele M (p)} = \frac{(2 \times 1787) + 3039}{2 \times 6129}$$

$$p = 0,54$$

$$q = 0,46$$

Frequenze attese all'equilibrio

$$p^2 = (0,54)^2 = 0,29$$

$$2pq = 2(0,54 \times 0,46) = 0,50$$

$$q^2 = (0,46)^2 = 0,21$$

Nel caso invece di malattie recessive come la Fibrosi Cistica invece:

- Sono distinguibili solo gli omozigoti affetti, con genotipo **aa**.
- Il principio di Hardy-Weinberg ci dice che la frequenza di **aa** dovrebbe essere q^2 .
- Per la fibrosi cistica nella popolazione europea, $q^2 = 1/2500$ (cioè la prevalenza della malattia tra i neonati).
- Per stimare q , usiamo la radice quadrata di entrambi i membri di questa equazione:
- $q = \sqrt{1/2500} = 1/50 = 0,02$.
- Poiché $p + q = 1$, $p = 0,98$.
- Qual è la frequenza degli eterozigoti, portatori sani della mutazione malattia?
- la frequenza degli eterozigoti è $2pq \approx 2q = 2/50 = 1/25$
 - $2pq = 2 \times 0,98 \times 0,02 = 0,039$
 - $2q = 1/25 = 0,04$

(L'A lo si indica come "p" ed è bene ricordare che per portatore si intende l'eterozigote) [NB: negli esercizi è importante tenere a mente che gli alleli sono il doppio degli individui].

Rh $\begin{cases} + = DD \text{ or } Dd \\ - = dd \end{cases}$

La stessa cosa si presenta per il gruppo sanguigno Rh^+ e Rh^- , il primo è dato da DD o da Dd, il secondo invece da dd.

Se conosciamo il numero di $Rh^- = 16\% = q^2$ possiamo fare la radice e ottenere $q = 4/10$

Possiamo quindi ricavare $p = 1 - q$. Ed infine l'eterozigosi $Dd = 2pq$

$$Rh(-) = 16\% = 16/100 = q^2$$

$$q = \sqrt{16/100} = 4/10 = 0,4$$

$$p = 1 - q = 0,6$$

$$Dd = 2pq = 2 \times 0,6 \times 0,4 = 0,48$$

Un altro esempio per vedere se la popolazione è in equilibrio:

Determinazione dello stato di equilibrio in una popolazione

Es. In caso di codominanza è invece possibile verificare se la popolazione sia effettivamente in equilibrio di HW

Dopo aver calcolato (per conta diretta) le frequenze alleliche

- si calcolano le frequenze genotipiche attese in caso di equilibrio e si confrontano con quelle osservate;

RISULTATI POSSIBILI

- ❖ le frequenze attese sono molto simili alle osservate: la popolazione è in equilibrio
- ❖ le frequenze attese si discostano molto dalle osservate: la popolazione **NON** è in equilibrio
- ❖ le frequenze attese si discostano dalle osservate ma non sappiamo valutare questa discrepanza: test del χ^2

[NB in caso di codominanza si può procedere direttamente per conta diretta].

Il χ^2 viene utilizzato in caso di lieve discrepanza per valutare se essa sia casuale o meno

I risultati del test statistico hanno valore esclusivo di probabilità. Vi sono possibili ipotesi e va valutato il livello di significatività stabilito dallo sperimentatore (a piacere) che corrisponde al valore P

H₀ (o «ipotesi nulla») = non esiste nessuna differenza tra i gruppi riguardo al parametro considerato. I gruppi sono fra loro uguali e la differenza osservata va attribuita al caso.

in base alle leggi della probabilità dobbiamo valutare in quale percentuale un risultato come quello da noi ottenuto, può essere dovuto a un puro caso.

Se ci accontentiamo di una precisione del 5% ($p < 0.05$) dobbiamo rilevare dalle apposite tabelle quale è il limite del valore osservato per confidare nella possibilità che il risultato da noi ottenuto sia dovuto al caso (in 5 valutazioni su 100)

Significativo al 5% = corro un rischio di 1/20 di essermi sbagliato a negare H₀

P < 0.01 = significativo all'1%

Viene arbitrariamente scelto il valore $P = 0,05$ come limite per accettare o rifiutare l'ipotesi.

Quando $P > 0,05$, si assume che la differenza sia dovuta al caso.

Quando $P < 0,05$, si presume che un fattore diverso dal caso abbia prodotto la differenza.

Il test indica la probabilità (P) che la differenza tra i valori osservati e quelli attesi sia dovuta al caso.

Il valore di χ^2 si calcola come segue:

$$\chi^2 = \sum \frac{(\text{numero osservato} - \text{numero atteso})^2}{(\text{numero atteso})}$$

Una volta calcolato χ^2 si deve cercare il corrispondente valore di Probabilità

Il numero di gradi di libertà (gdl) per l'equilibrio di HW è pari a:

$$[k(k-1)]/2$$

dove k è il no. di alleli

Sistema a 2 alleli	→	gdl = 1
Sistema a 3 alleli	→	gdl = 3
Sistema a 4 alleli	→	gdl = 6

Gradi di libertà = $(CF - 1) - (NA - 1) = CF - NA$
CF: numero delle classi fenotipiche
NA: numero degli alleli

Per ogni campionamento bisogna stabilire i gradi di libertà in base al numero di alleli stabilito seguendo la formula $[k(k-1)]/2$.

Esempio: [++ è omozigote polimorfa; +- sono eterozigoti; -- sono omozigoti che non presentano polimorfismo].

OSSERVATO	GENOTIPO			totale
	+/+	+/-	-/-	
Numero di individui	16	28	20	64
Numero di alleli +	32	28	0	60
Numero di alleli -	0	28	40	68
Somma degli alleli + e -	32	56	40	128

ATTESO: Frequenza allelica di + = $p = 60/128 = 0.469$
Frequenza allelica di - = $q = 68/128 = 0.531$

	p^2	$2pq$	q^2	
Frequenze genotipiche relative attese	0.220	0.498	0.282	1
Frequenze genotipiche assolute attese	14.1	31.9	18.0	64

Nota bene: le frequenze sono sempre frequenze assolute. Mai si possono utilizzare frequenze percentuali o relative. Le percentuali quindi vanno sempre ritrasformate in frequenze assolute moltiplicandole per il numero di osservazioni. Con i soli dati percentuali (senza il numero totale di osservazioni da cui ricavare le frequenze assolute) non si può fare alcun test.

[!Alert!: le frequenze assolute attese si ottengono moltiplicando le freq. Relative per il numero totale di individui; poi andranno confrontate con gli attesi

Applico quindi il test del chi quadro per verificare lo stato di equilibrio (bisogna capire il perché del lieve discostamento: caso o disequilibrio]

Verifica dello stato di equilibrio

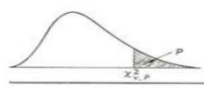
$$\chi^2 = \sum (\text{oss} - \text{att})^2 / \text{att}$$

$$\chi^2 = 0.955$$

GENO	OSSERV	ATTESI	(OSS-ATT) ²	(OSS-ATT) ² /ATT
++	16	14.1	3.61	0.256
+-	28	31.9	15.21	0.477
--	20	18	4	0.222
TOT	64	64	0	0.955

La differenza tra osservati e attesi non è significativa
La popolazione è in equilibrio

Gradi di libertà ? (numero delle classi - 1) - numero di parametri stimati
Gradi di libertà = (3-1) - 1 = 1



La funzione tabulata è $\chi^2_{n,p}$, valore superato con probabilità P in una distribuzione χ^2 con n gradi di libertà.

Gradi di libertà, n	Probabilità di un valore maggiore, P									
	0,975	0,900	0,750	0,500	0,250	0,100	0,050	0,025	0,010	0,001
1	—	0,02	0,10	0,45	1,32	2,71	3,84	5,02	6,63	10,83
2	0,05	0,21	0,58	1,39	2,77	4,61	5,99	7,38	9,21	13,82
3	0,22	0,58	1,21	2,37	4,11	6,25	7,81	9,35	11,34	16,27
4	0,48	1,06	1,92	3,36	5,39	7,78	9,49	11,14	13,28	18,47
5	0,83	1,61	2,67	4,35	6,63	9,24	11,07	12,83	15,09	20,52
6	1,24	2,20	3,45	5,35	7,84	10,64	12,59	14,45	16,81	22,46
7	1,69	2,83	4,25	6,35	9,04	12,02	14,07	16,01	18,48	24,32
8	2,18	3,49	5,07	7,34	10,22	13,36	15,51	17,53	20,09	26,12
9	2,70	4,17	5,90	8,34	11,39	14,68	16,92	19,02	21,67	27,88
10	3,25	4,87	6,74	9,34	12,55	15,99	18,31	20,48	23,21	29,59
11	3,82	5,58	7,58	10,34	13,70	17,28	19,68	21,92	24,72	31,26
12	4,40	6,30	8,44	11,34	14,85	18,55	21,03	23,34	26,22	32,91
13	5,01	7,04	9,30	12,34	15,98	19,81	22,36	24,74	27,69	34,53
14	5,63	7,79	10,17	13,34	17,12	21,06	23,68	26,12	29,14	36,12
15	6,27	8,55	11,04	14,34	18,25	22,31	25,00	27,49	30,58	37,70

non significativo

$P < 0.01$
molto significativo

Si è potuto stabilire che la popolazione fosse in equilibrio grazie al confronto con valori indicati sulla tabella, corrispondenti al determinato grado di libertà (1 in questo specifico caso) e si è potuto notare che il valore (compreso tra i due cerchiati in rosso) rientrasse nel range.

Quando invece si trova una differenza statisticamente significativa, bisogna porsi alcune domande:

- Il campione è rappresentativo (almeno 100 persone minimo)?
- La tecnica utilizzata è corretta? Per ogni campione vanno infatti usati specifici test e potrebbero esserci errori di valutazione.

Non è quindi detto che, automaticamente, la popolazione non sia in equilibrio.

FINE DELLA SBOBINA

!DI SEGUITO DEGLI ESERCIZI SVOLTI IN CALSSE CON RISPETTIVI RISULTATI!

ESEMPIO N°1 :

- Stai studiando una popolazione in cui la frequenza di individui con un genotipo omozigote recessivo è dell'1%. Supponendo che la popolazione sia in equilibrio di Hardy-Weinberg, calcola:
 - La frequenza dell'allele recessivo
 - La frequenza dell'allele dominante
 - La frequenza del genotipo eterozigote
 - La frequenza del genotipo omozigote dominante
 - La frequenza del fenotipo dominante

uso **HW** perché si tratta di una malattia recessiva in cui non è distinguibile l'eterozigote dall'omozigote (sano) altrimenti avrei fatto la conta diretta

$$p^2 + 2pq + q^2 = (p + q)^2 = 1$$

AA Aa aa

- $q = \sqrt{0,01} = 0,1$
- $p = 1 - q = 0,9$
- $2pq = 2 \times 0,9 \times 0,1 = 0,18$
- $p^2 = 0,9 \times 0,9 = 0,81$
- $p^2 + 2pq = 0,81 + 0,18 = 0,99$

ESEMPIO N°2 :

malattia autosomica recessiva:

Determinazione della *v* genica e del portatore per una malattia AR

Esempio 1: La fenilchetonuria è una malattia autosomica recessiva la cui frequenza alla nascita è circa 1/10.000

FENOTIPI	AA	Aa	aa
	normale	portatore normale	affetto

	p^2	$2pq$	q^2
--	-------	-------	-------

FREQUENZA DEGLI AFFETTI ALLA NASCITA	$q^2 = 1/10000$
FREQUENZA GENICA DEL GENE DELETERIO	$q = 1/100 = 0,01$
FREQUENZA GENICA DEL GENE NORMALE	$p = 99/100 = 0,99$
FREQUENZA DEGLI INDIVIDUI PORTATORI	$2pq = 2 \times 0,01 \times 0,99 = 0,02$

Determinazione della *v* genica e del **CARRIER** sani per una malattia AR

Calcolo frequenza dell'insieme degli alleli CF (AR)

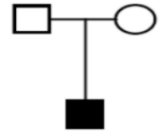
$$\begin{aligned} \text{Freq Malati } (q^2) &= 1/2500 = 4 \times 10^{-4} \\ \text{Freq allele } (q) &= \sqrt{q} (4 \times 10^{-4}) = 0.02 \text{ (o } 1/50) \end{aligned}$$

Diventa possibile calcolare quanti sono i carrier sani (= eterozigoti CF/nonCF), infatti
 $p = 1 - q = 1 - 1/50 = 49/50 (=0.98 \approx 1)$, quindi:

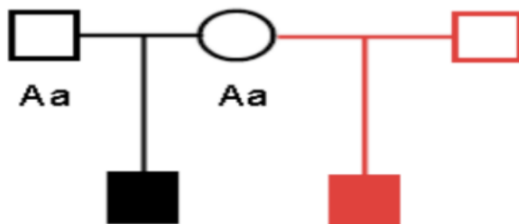
$$2pq = 2 \times 49/50 \times 1/50 = 1/25$$

Probabilità che due eterozigoti si incrocino è:
 $1/25 \times 1/25 = 1/625$

Tra i loro figli 1/4 avrà la CF, quindi la frequenza totale sarà: $1/625 \times 1/4 = 1/2500$
 Cioè come osservato



**Se il genitore di un bambino con CF si risposa,
 qual è il rischio di avere un figlio malato?**



$$\begin{array}{ccc} \text{Rischio} & & \text{Rischio} \\ \text{portatore} & \times & \text{portatore} \\ \\ 1 & \times & 1/25 \end{array}$$

$$1/25 \times 1/4 = 1/100$$

Il principio di HW è importante anche applicato alle malattie legate al cromosoma X

Principio di Hardy-Weinberg: geni legati all'X

I geni legati all'X sono un caso particolare perché i maschi hanno un solo cromosoma X
 Frequenze genotipiche tra le femmine: $HH = p^2$; $Hh = 2pq$; $hh = q^2$

Le frequenze genotipiche tra i maschi sono le stesse delle frequenze alleliche:
 $H = p$, $h = q$

TABLE 7-5			
X-Linked Genes and Genotype Frequencies (Color-Blindness)			
Sex	Genotype	Phenotype	Incidence (Approximate)
Male	X^+	Normal color vision	$p = 0.92$
	X^b	Color blind	$q = 0.08$
Female	X^+/X^+	Normal (homozygote)	$p^2 = (0.92)^2 = 0.8464$
	X^+/X^b	Normal (heterozygote)	$2pq = 2(0.92)(0.08) = 0.1472$
		Normal (total)	$p^2 + 2pq = 0.9936$
	X^b/X^b	Color blind	$q^2 = (0.08)^2 = 0.0064$

ESEMPIO N°3:

Determinazione degli eterozigoti per una malattia X-linked

Daltonismo comune X-linked R
In Svizzera colpisce 1 su 12 maschi

In realtà le donne daltoniche sono di meno ...2
geni su Xq28: protanopia e deutanopia
OMIM 303900

Quali sono le frequenze nelle donne portatrici e affette?

	maschi		femmine		
Genotipi	X^A	X^a	$X^A X^A$	$X^A X^a$	$X^a X^a$
Frequenze	p	q (1/12) (0.08)	p^2	$2pq$	q^2

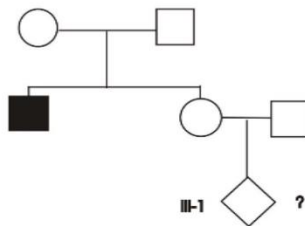
Se $q=1/12$ ($X^a Y$) $p=1-q = 11/12$ ($X^A Y$) (0.92)

Donne eterozigoti= $2 \times 1/12 \times 11/12 = 22/144$ (15%)
Donne daltoniche= $q^2 = 1/144$ (0,7%)

	Maschi		Femmine		
Genotipi	$X(A)$	$X(a)$	$X(A)X(A)$	$X(A)X(a)$	$X(a)X(a)$
Frequenze	p	q	p^2	$2pq$	q^2
osservate		0,08			
calcolate	0,92		0,846	0,147	0,007

ESEMPIO N°4 :

- Una malattia autosomica recessiva ha una frequenza di 1/3.600. Se la sorella di un paziente sposa un individuo della stessa popolazione, che è fenotipicamente sano e non ha parenti malati in famiglia per quella malattia, qual è la probabilità che questa coppia generi un figlio affetto?



$$q^2 = 1/3,600$$

$$q = \sqrt{1/3,600} = 1/60 \quad p = 59/60 \approx 1$$

$$2pq \approx 2q = 2 \times 1/60 \approx 1/30$$

$$R = 2/3 \times 1/30 \times 1/4$$

ESEMPIO N° 5 :

La frequenza dei maschi affetti da una malattia recessiva legata all'X è di 1/5.000.

Quale sarà la frequenza di portatori sani nella stessa popolazione?

$$q = 1/5,000 \quad p \approx 1$$

$$2pq = 2q = 2/5,000 = 1/2,500$$

1. Consideriamo una popolazione in cui sono presenti i seguenti genotipi con le frequenze indicate: AA (0.04), Aa (0.32), aa (0.64). Qual è la frequenza dell'allele A?
- a) 0.36
 - ☒ b) 0.20
 - c) 0.68
 - d) 0.04
2. In un campione di 50 individui si osservano i seguenti genotipi per un dato locus : 1/1= 13, 1/2= 34, 2/2= 3; quali sono le frequenze alleliche di 1 e 2?
- a) 0,5; 0,5
 - b) 0,13; 0,37
 - c) 0,27; 0,73
 - ☒ d) 0,6; 0,4
5. In una certa popolazione in equilibrio un individuo su dieci è emofilico. Qual è la frequenza delle donne eterozigoti?
- a) 0.5
 - ☒ b) 0.2
 - c) 0.9
 - d) 0.81