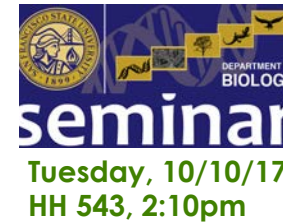


Class 18 10/4/17

Chi square test/human genetics

- Announcements
- Class administration
- Check iLearn for suggested problems
- **Monday 10/9 – Alternate Activity will be available on iLearn**
- **Wednesday 10/11 – Exam 2 (Classes 11 – 19)**
- **Office hours:**
 - Extra office hours on TUE 10/10 4:30-6pm, HH525

1



Biol 572/872 Ecology, Evolution, & Conservation Biology Colloquium

<http://biology.sfsu.edu/content/EEC>

TBA

2



Romberg Tiburon Center Seminar Series

<http://rtc.sfsu.edu/seminar/index.htm>

Wednesday, 10/4/17
Bay Conference Center, 3:30PM



Ben Becker
Point Reyes National Seashore
*Climate Change or Fisheries:
What is Driving Long-term
Trophic Level Decline in
California Current Seabirds?*



Biol 871 Colloquium in Microbiology, Cell & Molecular Biology

<http://biology.sfsu.edu/content/MCMB>



Katja Brückner
UCSF
*How sensory neurons drive organ
development and homeostasis –
Drosophila models for
hematopoiesis, immunity and
epithelial plasticity*

<https://bruecknerlab.ucsf.edu/welcome-br%C3%BCckner-lab>

4

3

Process of using chi-square analysis to test goodness of fit

- Establish a null hypothesis: 1:1, 3:1, etc.
- Plug data into the chi-square (χ^2) formula.
- Determine if null hypothesis is either (a) rejected or (b) not rejected.
- If rejected, propose alternate hypothesis.
- Chi-square (χ^2) analysis factors in
 - deviation from expected result and
 - sample size to give measure of the **goodness of fit** of the data

5

Chi-square formula

$$\chi^2 = \sum \frac{(o - e)^2}{e}$$

where

- o = observed value for a given category,
- e = expected value for a given category, and
- sigma (Σ) is the sum of the calculated values for each category of the ratio
- Once χ^2 is determined, it is converted to a probability value (p) using the **degrees of freedom** (df) = $n - 1$, where n = the number of different categories for the outcome

6

Chi-square - Example 1

Phenotype	Expected	Observed
A	750	760
a	250	240
total	1000	1000

- Null Hypothesis: Data fit a 3:1 phenotypic ratio.

$$\chi^2 = \sum \left[\frac{(o - e)^2}{e} \right] = \left[\frac{(760 - 750)^2}{750} + \frac{(240 - 250)^2}{250} \right]$$

$$\chi^2 = 0.53$$

- degrees of freedom** = (number of categories - 1) = $2 - 1 = 1$
- Use Table 4-1 to determine p - on next slide

7

Is the difference between observed and expected due to chance?

Table 4-1 Critical Values of the χ^2 Distribution

df	P									df
	0.995	0.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005	
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	12
13	3.565	5.009	7.033	12.340	19.812	22.362	24.736	27.688	29.819	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	15

Likely due to chance
Fails to reject null hypothesis

Unlikely due to chance
Reject null hypothesis

Degrees of freedom = 1

$$\chi^2 = 0.53$$

$$0.50 > p > 0.10$$

likely

unlikely

8

Interpretation of p

- 0.05 is a commonly-accepted cut-off point.
- $p > 0.05$ means that the probability is greater than 5% that the *observed deviation is due to chance alone*; therefore **we fail to reject the null hypothesis**.
- $p < 0.05$ means that the probability is less than 5% that observed deviation is due to chance alone; therefore **null hypothesis is rejected**. Reassess assumptions, propose a new hypothesis.

9

Conclusions:

- χ^2 less than the **critical value 3.84** ($p > 0.05$) means that we fail to reject the Null Hypothesis (3:1 ratio).
- In our example, $p = 0.48$ ($p > 0.05$) means that we **fail to reject the Null Hypothesis** (3:1 ratio).
- When you solve this type of problem, you need to STATE YOUR CONCLUSION:
- "The chi-square analysis shows that $p > 0.05$; therefore, **we fail to reject our hypothesis**: The inheritance of A and a phenotypes is the result of a monohybrid cross, where A is dominant to a, because the F2 data fit the expected 3:1 ratio."

10

χ^2 - Example 2

- In a set of fruit fly crosses, I observed the following F2 data:
792 long-winged (wildtype) flies (**Phenotype D**)
208 dumpy-winged flies (**Phenotype d**)
- **My null hypothesis:** "dumpy wing is inherited as a Mendelian recessive trait ($D > d$), and the F2 phenotypic ratios are consistent with a 3D:1d ratio (*there is no real difference between the expected values and the observed values - the apparent difference is due to chance*)"

11

Clicker Question

- How many of the **dominant phenotype** do you expect to observe?

Expected Ratio	Observed (o)	Expected (e)	Deviation	d^2	d^2/e
3/4	792	A. 792			
		B. 208			
		C. 750			
		D. 250			
1/4	208	E. Not able to calculate			
	1000 total progeny				

12

χ^2 analysis

- Degrees of freedom = 2 - 1 = 1

Expected Ratio	Observed (o)	Expected (e)	Deviation d = o - e	d ²	d ² /e
3/4	792	750	42	1764	2.352
1/4	208	250	-42	1764	7.056
	1000 total progeny			$\chi^2 =$	9.408

13

Clicker Question: Evaluate the deviation

Table 4-1 Critical Values of the χ^2 Distribution

	P									
df	0.995	0.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005	d
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	A. R B. F C. N	
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666		
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209		
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725		
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	12
13	3.565	5.009	7.041	12.340	19.812	22.362	24.736	27.688	29.819	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	15

- $\chi^2 = 9.408$
- degrees of freedom = 1

Is the difference between observed and expected due to chance?

- A. Reject
B. Fail to Reject
C. Not enough info to evaluate

14

Is the difference between observed and expected due to chance?

Table 4-1 Critical Values of the χ^2 Distribution

	P									
df	0.995	0.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005	df
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	12
13	3.565	5.009	7.041	12.340	19.812	22.362	24.736	27.688	29.819	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	15

- $\chi^2 = 9.408$
- degrees of freedom = 1

$$.01 > p > .001$$

$$p < 0.05$$

Null hypothesis is rejected

Dumpy is NOT inherited as a recessive trait

15

Another way to state this...

- P-value \sim 0.001 means that Chi-square values equal to or greater than 9.408 are expected to occur 0.1% of the time due to random chance alone; that is, when the null hypothesis is true.
 - Low probability of this high chi-square value
 - Reject the null hypothesis

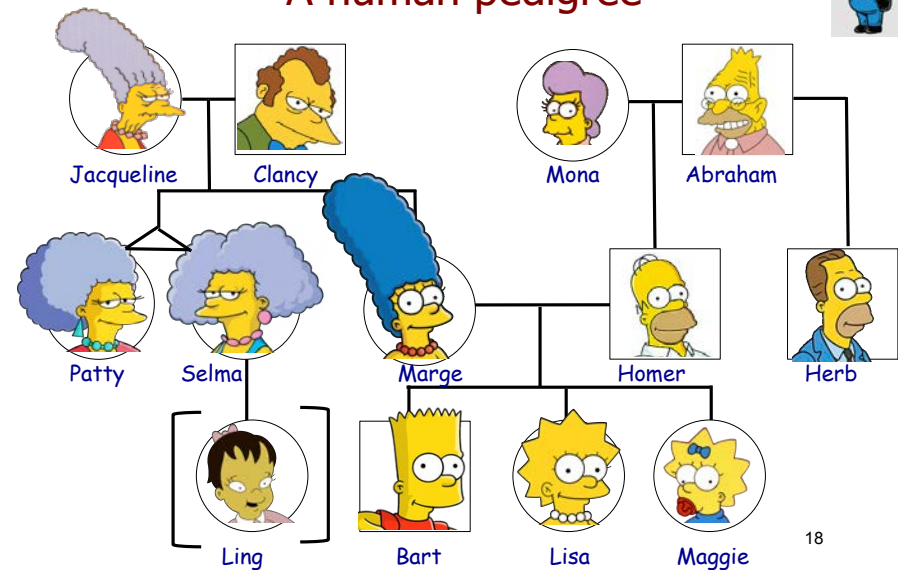
16

Inheritance of human traits

- How do we predict the mode of inheritance in humans...?
 - Experimental constraints
 - Cannot design matings
 - Do not recover hundreds of progeny
- A **pedigree** shows a family tree with respect to a given trait.
- Pedigree analysis reveals patterns of inheritance.

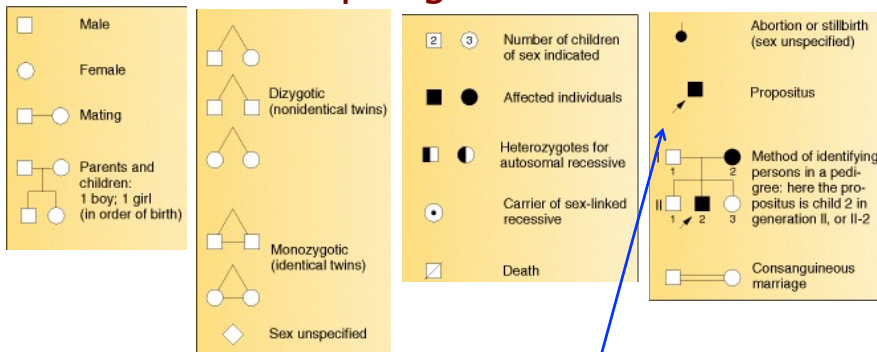
17

A human pedigree



18

Standard pedigree conventions



Or "proband"
Or "index case"

See next slide for symbols for pregnancies, and pregnancies not carried to term

19
Figure 2-16

Standard pedigree symbols - 2

	Male	Female	Sex unknown
Pregnancy (P)			
Spontaneous abortion (SAB), ectopic (ECT)			
Affected SAB			
Termination of pregnancy (TOP)			
Affected TOP			

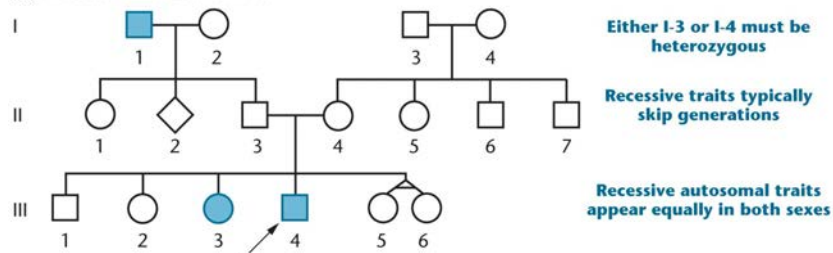
Standardized pedigree symbols and nomenclature for pregnancies and pregnancies not carried to term

<http://what-when-how.com/genetics/genetic-family-history-pedigree-analysis-and-risk-assessment/>

20

Autosomal recessive trait

(a) Autosomal Recessive Trait

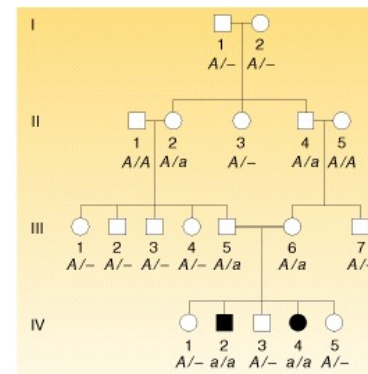


- Autosomal recessive traits can skip generations (appear in progeny of unaffected persons) and affect both males and females equally.

21

Klug, 10th edition, Figure 3.13

Autosomal recessive trait



- Autosomal recessive traits can skip generations (appear in progeny of unaffected persons) and affect both males and females equally.
- Rare recessive traits will often be expressed in consanguineous (blood relations) matings (first cousins)
- Example: The albino (aa) mutation inactivates the gene for tyrosinase enzyme, which normally converts tyrosine to melanin in the skin, hair and eyes.
 - Non-albino is AA or Aa

Figure 2-17