Applied Epistemology:

A Framework for how we know things scientifically.

A Refresher on Learning Path I: Hypothesis Testing

Agenda:

1. H_O/H_A: Our model of the test universe (the distribution of the variable)
2. Test & assumptions: are the assumptions met? Is the test valid?
3. Quantitative evidence: **p-value**, or critical value.
   - False positive =Type I (α), False Negative = Type II (β), Type III errors
   - Sensitivity, Specificity, Power → confusion matrix, ROC/AUC curve
   - Confusion Matrix
4. Conclusion & uncertainty/estimation
5. **Z-scores, χ^2 Goodness-of-fit test**, and **χ^2 Contingency test**

Perspectivism – why assumptions are important: https://hdsr.mitpress.mit.edu/pub/qasl4fza/release/3

Errors in hypothesis testing:

Type I (α) = False Positive=0.05

P[type I]=P[rejecting Ho|Ho is true]

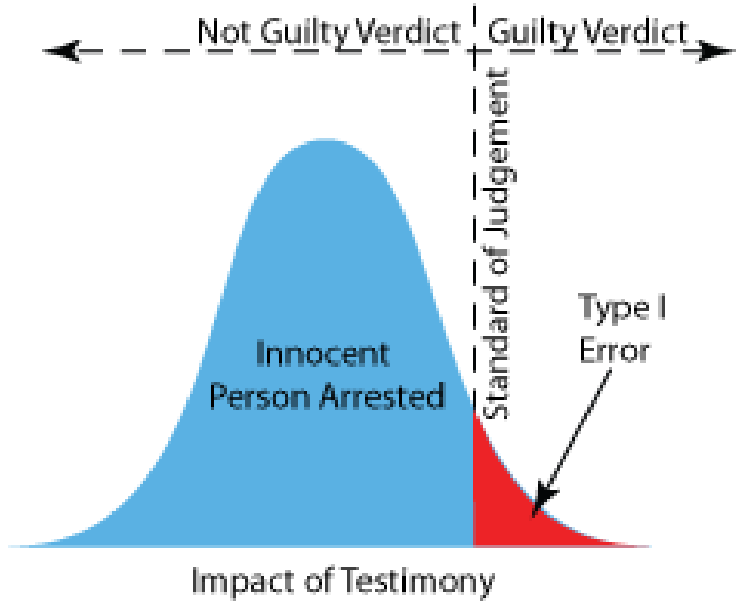
Type II (β) = False Negative

P[type II]=P[Fail-to-reject Ho|Ho is not true]

Type I (α) error:

Rejecting a true null hypothesis

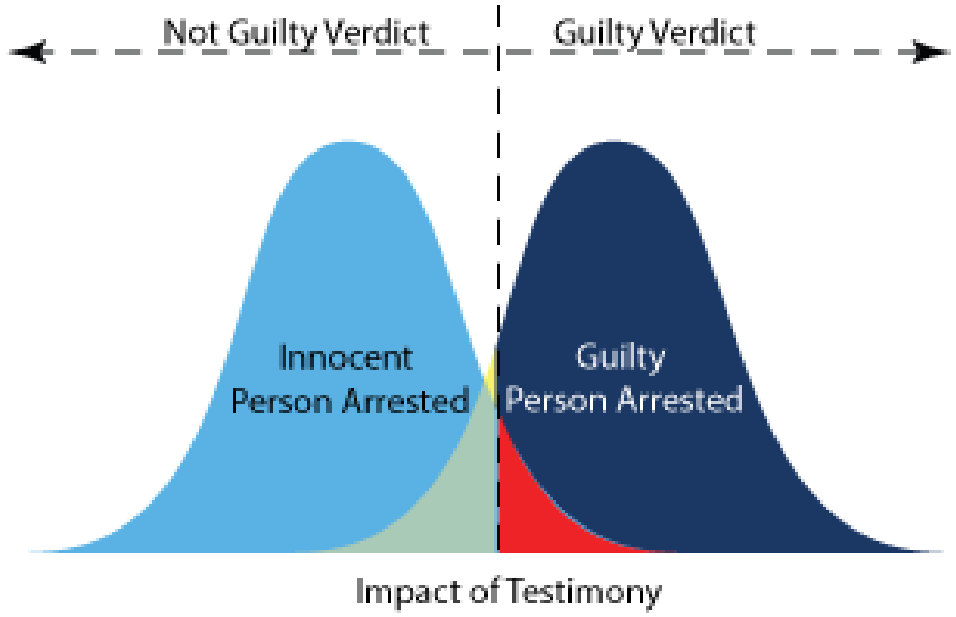
$P(\text{reject } H_0 | H_0 = \text{true}) = \alpha$



Type II (β) error:

Not rejecting a false null hypothesis

$P(\text{Fail to reject } H_0 | \text{Ho=Not True}) = \beta$



http://www.intuitor.com/statistics/T1T2Errors.html

https://wise.cgu.edu/wise-tutorials/tutorial-hypothesis-testing/type-1-error-rejecting-a-true-null-hypothesis/

Visualize Type I/II errors: One-sample Test of Means (Z test)

Choose Tail of the Test

- ○ One Tail, Upper Tail
- ○ One Tail, Lower Tail
- ◉ Two Tail

Choose Plot to Display

- ☑ Show Null Hypothesis Sampling Distribution
- ☑ Show Alternative Hypothesis Sampling Distribution

Choose alpha control via slider or menu

- ○ Choose among several fixed alpha levels
- ◉ Use a slider for alpha choice

Alpha, Type I Error rate

0.0005 | 0.05 | 0.15

0.0005 0.0155 0.0305 0.0455 0.0605 0.0755 0.0905 0.1055 0.1205 0.1355 0.15

Null Hypothesis Mean

100

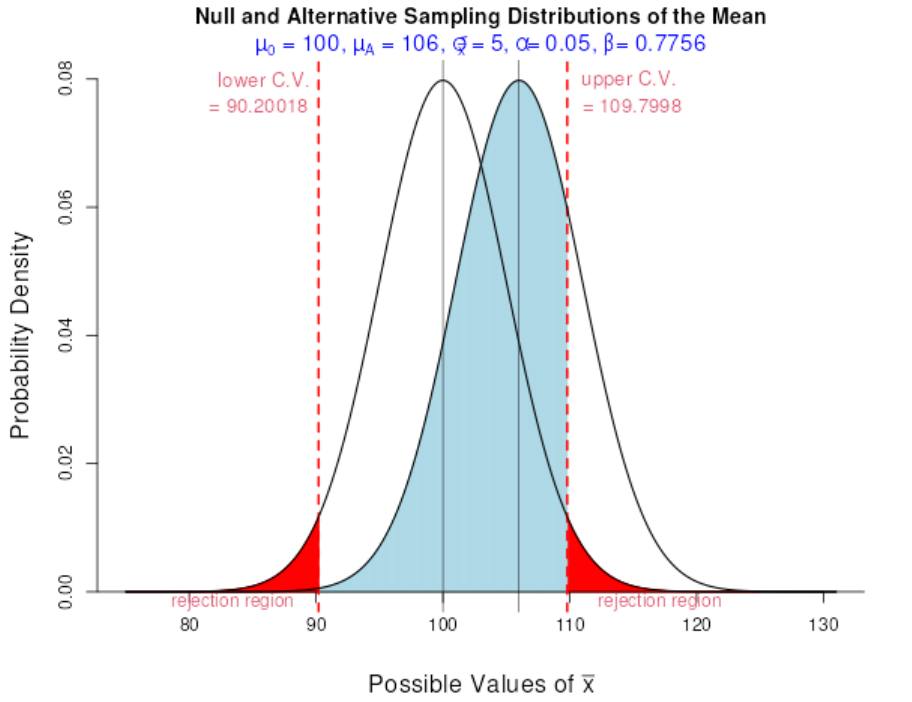
Alternative Hypothesis Mean

106

Standard Error of the Mean

5

Plot | About



Download plot as PDF

Download plot as png

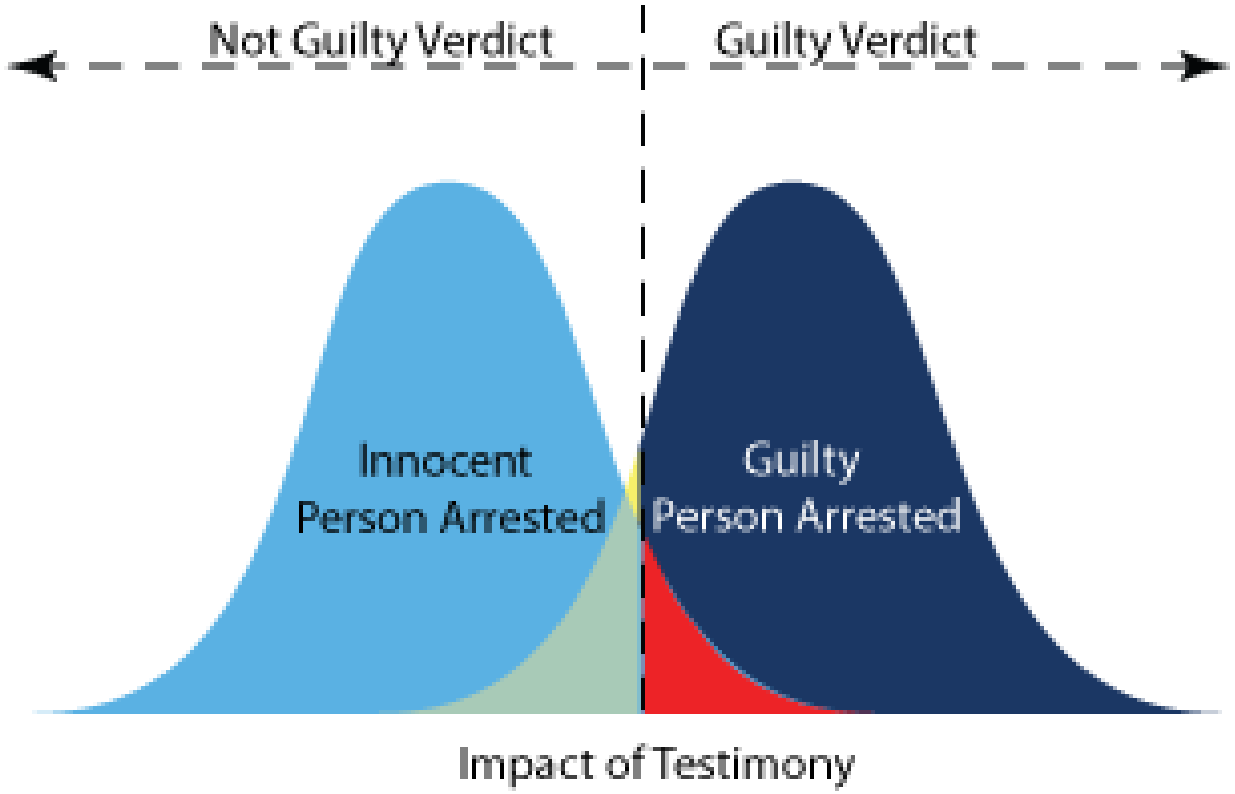
https://shiny.rit.albany.edu/stat/betaprob/

	No Disease (H_0 true)	Disease (Ho is not true; H_A true)
Fail To Reject H_0	**No Error** P[FTR|Ho is true] (True Negative)	**Type II** P[FTR|Ho is not true] (False Negative)
Reject H_0	**Type I** P[reject|Ho] (False Positive)	**No Error** Power P[Reject|Ho is not true] (True Positive)

Specificity = $\frac{TN}{TN+FP}$

Sensitivity = $\frac{TP}{TP+FN}$

https://en.wikipedia.org/wiki/Sensitivity_and_specificity



Generally, type I errors are the ones that we are concerned with in biology.
Although, there are circumstances when we are more concerned with type II errors (i.e.) Medicine

There is a trade-off between type I error and type II error

Add more data points, n, and you are able to discriminate between smaller differences in the null and alternate hypotheses!

Power is the ability of a test to reject a false null hypothesis

Power = 1 - P(FTR H_0| H_A)= P(Reject H_0| H_A)

Power = 1-β

Sensitivity = 1-Type II = TP / (TP+FN)

Specificity = 1-Type I = TN / (TN+FP)

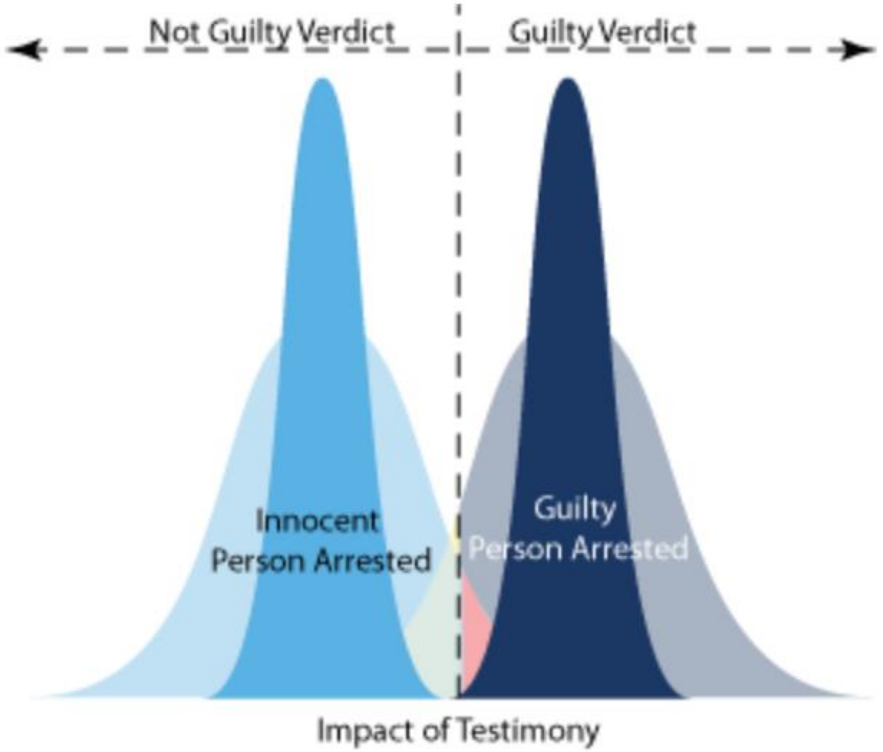


figure 5. The effects of increasing sample size or in other words, number of independent witnesses.

http://www.intuitor.com/statistics/T1T2Errors.html

Two clinical trials are carried out which both test the same null hypothesis under the same conditions with α = 0.05. Trial **A** has **45 individuals** and Trial **B** has **100 individuals**. Power=1-type II (Beta)

Which of the following is true about the two trials described above:

a. Study A has higher probability of type I error than Study B and Study B has a higher probability of type II error than Study A

b. Study A has a lower probability of type I error than Study B and Study B has a lower probability of type II error than Study A

c. Study A has the same type I error as Study B and Study A has a higher probability of type II error than Study B.

d. Study A has the same type I error as Study B and Study B also has a higher probability of type II error than Study A

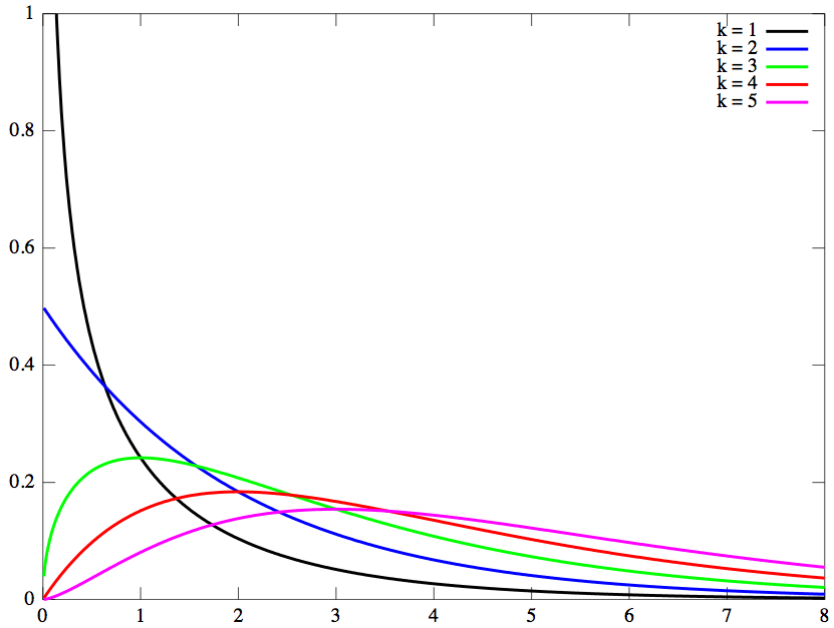
With multiple simultaneous hypothesis tests, you also need to ensure that you haven't **inflated alpha!** (see the horoscope and illness example from Biostatistics I)

You can use the Hypothesis Testing Framework with any sample distribution, not just the Normal Distribution

H_0: *The data come from a particular discrete probability distribution*
H_A: *The data do not come from that distribution*

Step 2: $\chi 2$ Goodness of fit test (or contingency test):

$$\chi^2_{df} = \sum_i \frac{(Observed_i - Expected_i)^2}{Expected_i}$$



Are assumptions met?

H_0: *The data come from a particular discrete probability distribution*

H_A: *The data do not come from that distribution*

Step 2: $\chi 2$ Goodness of fit test (or contingency test):

$$\chi^2_{df} = \sum_i \frac{(Observed_i - Expected_i)^2}{Expected_i}$$

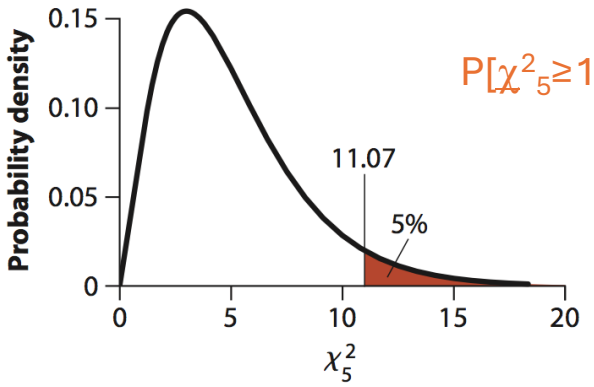
Step 3:

DF= # categories – 1- # estimated parameters*

Finding **critical values for χ^2** distribution:

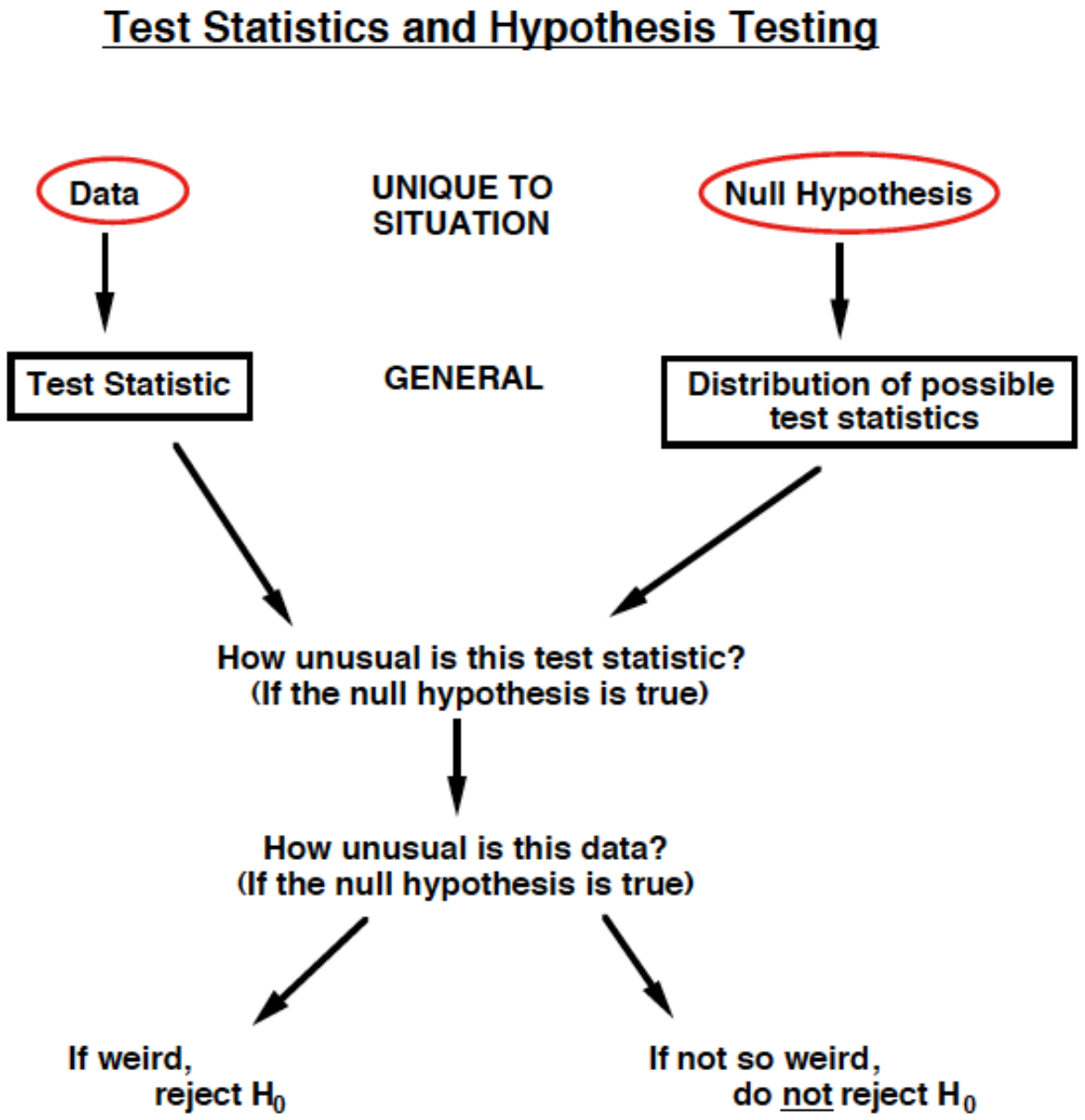
Statistical Table: https://www.math.arizona.edu/~jwatkins/chi-square-table.pdf

df	0.995	....	0.950	0.05
1	0.000		0.00393	3.841
...				
5	0.412		1.145	**11.070**
6	0.676		1.635	12.592



$P[\chi^2_5 \geq 11.070] = 0.05$

* df are calculated differently for contingency test



"A model is a mathematical tool that mimics how we *think* a natural process works.."

Life is interesting when a model doesn't fit the data because it suggests that at least one of the major assumption about how we think about the process is wrong

All models are wrong, but some are useful
- George E.P. Box

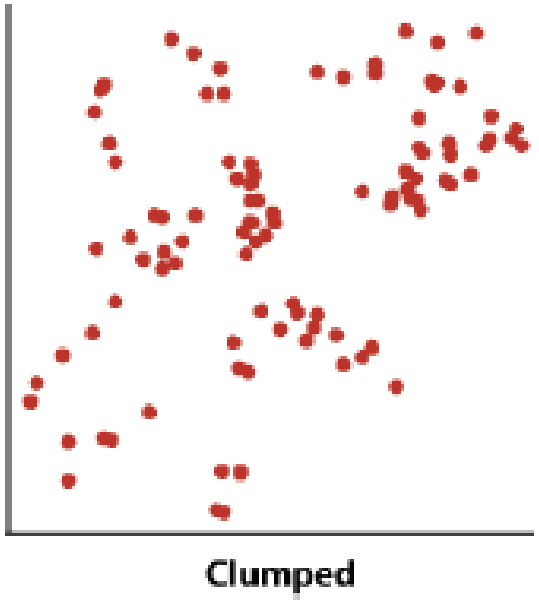
You're testing whether **incidence of infection** differs between **vaccinated** and **unvaccinated** groups.

Which test applies: binomial test, χ^2 Contingency, χ^2 GOF, Fisher's exact? What would be the null sampling distribution?

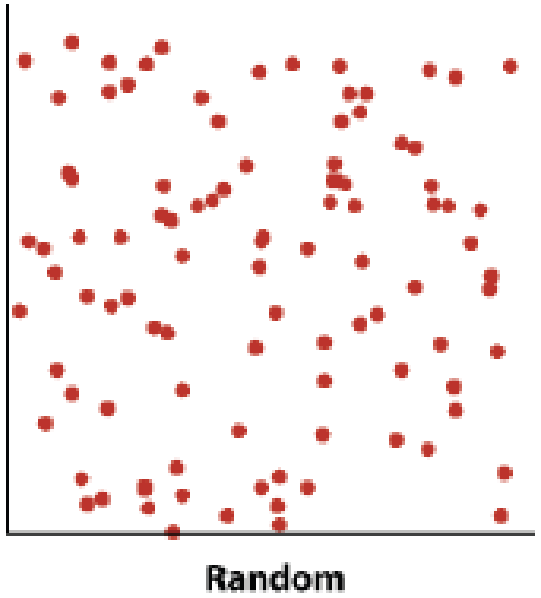
Why?

Fitting the Poisson Distribution:

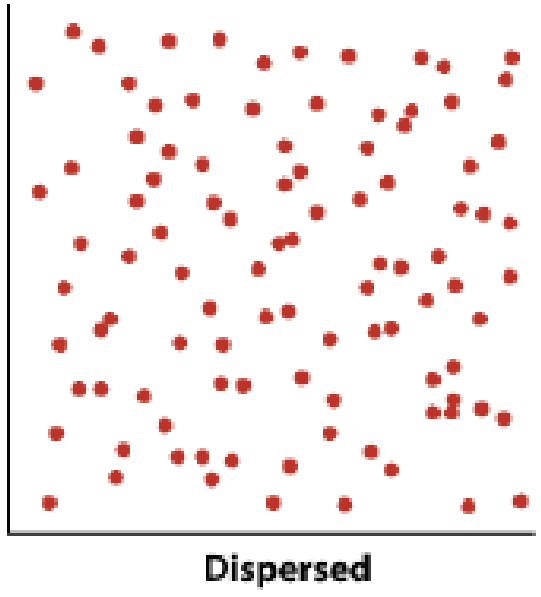
The Poisson Distribution describes the probability of getting X successes in a block of time or space when the successes happen independently of each other and occur with equal probability at every point in time or space.



Variance > mean



Variance = mean



Variance < mean

Poisson Distribution:

$$P[X] = \frac{e^{-\mu}\mu^{X}}{X!}$$

Example: Mutations are rare events, and we typically model mutations as having the same rate at each position in the genome and that each position in the genome is independent of every other position.

Is the following 100 replicates of spontaneous mutations in HIV consistent with a Poisson distribution of mutations?

Number of Mutations	Observed Frequency
0	1
1	5
2	12
3	10
4	14
5	20
6	19
7	10
8	6
9	1
10	2

Step 1

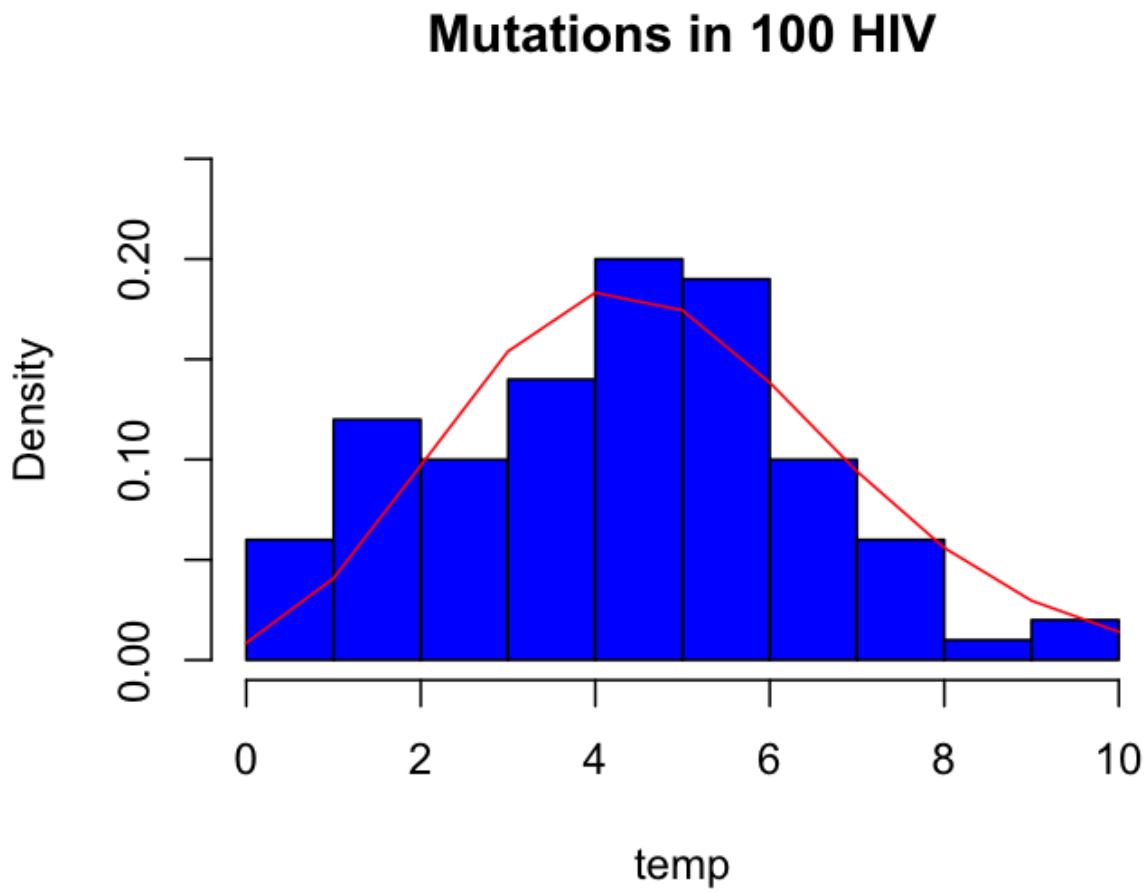
H_0: The number mutations per nucleotide per generation has a Poisson distribution

H_A: The number of mutations per nucleotide per generation does *NOT* have a Poisson distribution

Step 2

Estimate μ: $\bar{X} = \frac{1X0+5X1+12X2+10X3+14X4+20X5+19X6+10X7+6X8+1X9+2X10}{100} = 4.76$

$s^2 = 4.40$



Number of Mutations	Observed Frequency	Expected Frequency
0	1	
1	5	
2	12	
3	10	
4	14	
5	20	
6	19	
7	10	
8	6	
9	1	
10	2	

$$P[X] = \frac{e^{-\mu}\mu^{X}}{X!}$$

Number of Mutations	Observed Frequency	Expected Frequency
0	1	0.00854*100=0.85
1	5	0.0406*100=4.06
2	12	0.0966*100=9.66
3	10	0.153*100=15.3
4	14	0.183*100 = 18.2
5	20	0.173*100=17.3
6	19	0.137*100=13.7
7	10	0.093*100=9.3
8	6	0.055*100=5.5
9	1	0.029*100=2.9
10	2	0.0138*100= 1.38

These don’t quite add up to 100, due to rounding

Number of Mutations	Observed Frequency	Expected Frequency
0	1	0.00854*100=0.85
1	5	0.0406*100=4.06
2	12	0.0966*100=9.66
3	10	0.153*100=15.3
4	14	0.183*100 = 18.2
5	20	0.173*100=17.3
6	19	0.137*100=13.7
7	10	0.093*100=9.3
8	6	0.055*100=5.5
9	1	0.029*100=2.9
10	2	0.0138*100= 1.38

These don't quite add up to 100, due to rounding

Number of Mutations	Observed Frequency	Expected Frequency
0 +1	6	4.06+0.85=4.91
2	12	0.0966*100=9.66
3	10	0.153*100=15.3
4	14	0.183*100 = 18.2
5	20	0.173*100=17.3
6	19	0.137*100=13.7
7	10	0.097*100=9.3
8+	9	8.4+1.4 = 9.8

We had to combine bins in order to met the assumptions:

≥80% *expected* cells must be ≥ 5; no expected cell can have a value <1.0.

This will impact our degrees of freedom calculation!

Step 3:

Calculate your test statistic

$$\chi^2 = \frac{(4.91-6)^2}{4.91} + \frac{(9.66-12)^2}{9.66} + \frac{(15.30-10)^2}{15.30} + \frac{(18.2-14)^2}{18.2} + \frac{(17.3-20)^2}{17.3} + \frac{(13.70-19)^2}{13.70} + \frac{(9.3-10)^2}{9.3} + \frac{(9.8-9)^2}{9.8}$$

$$= 6.20$$

Step 3:

dof = 8 categories **– 1 – 1 estimate = 6** degrees of freedom

Step 4:

Critical value for χ^2 is given in statistical table found at:
https://www.math.arizona.edu/~jwatkins/chi-square-table.pdf
For alpha=0.05, df=6: 12.592
Therefore, we FTR the Poisson distribution. **BUT THERE IS MORE WE CAN SAY….**

Variance = Mean:

If Variance > Mean, then CLUMPED

- visual hint: histogram is 'u-shaped'

If Variance < Mean, then DISPERSED

- points are spread uniformly in space or time

- This may be a bit confusing if you are familiar with molecular genetics, because we refer to the "over dispersed molecular clock" which is really saying that variance > mean number of substitutions. Sometimes, terminology is ambiguous!