We conduct an experiment to see if a coin is fare. We flip the coin twice and record the number of heads. After repeating this 100 times these are our results:

0H	1H	2H
28	51	21

We wish to test the hypothesis that this is a fair coin, to do that we need

- Expected Probabilities: $P(0H) = P(T \text{ and } T) = P(T) \cdot P(T) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

  since the two events are independent

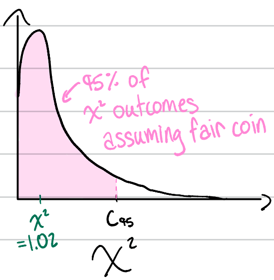
  $P(2H) = P(H \text{ and } H) = P(H) \cdot P(H) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

  $P(1H) = P([H \text{ and } T] \text{ or } [T \text{ and } H]) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
- Expected counts: Total flips · Probability

  $0H = 100 \cdot \frac{1}{4} = 25, \quad 1H = 100 \cdot \frac{1}{2} = 50, \quad 2H = 100 \cdot \frac{1}{4} = 25$
- What would we do to test? Chi-squared Goodness of Fit Test!

$$\chi^2 = \sum \frac{(\text{Exp} - \text{Observed})^2}{\text{Exp}} = \frac{(25-20)^2}{25} + \frac{(50-51)^2}{50} + \frac{(25-21)^2}{25} = \underline{1.02}$$

df in this case is 2 since we had 3 total outcomes so $C_{95} = 5.99$



$C_{95} = 5.99 > 1.02 = \chi^2$

so we fail to reject

What if we want to check if birth rate and stork population in a county are correlated.
In 200 counties we observe:

Birth rate $\begin{cases} 100 \text{ Low birth rate} \\ 100 \text{ High birth rate} \end{cases}$ $P(LB)=0.5$, $P(HB)=0.5$

Stork pop $\begin{cases} 150 \text{ Low storks} \\ 50 \text{ High storks} \end{cases}$ $P(LS)=0.75$, $P(HS)=0.25$

- Expected probability: We are expecting independence so...

$P(LB \text{ and } LS) = P(LB)\cdot P(LS) = \frac{1}{2}\cdot\frac{3}{4} = 3/8$

$P(HB \text{ and } HS) = P(HB)\cdot P(HS) = \frac{1}{2}\cdot\frac{1}{4} = 1/8$

$P(HB \text{ and } LS) = 3/8$

$P(LB \text{ and } HS) = 1/8$

- Expected Counts: Total count · probability

$3/8\cdot 200$ $\quad$ $1/8\cdot 200$

	LS	HS	
LB	75	25	100
HB	75	25	100
	150	50	

Observed:

	LS	HS	
LB	90	10	100
HB	60	40	100
	150	50	

- What would we do to test? Chi-squared Goodness of Fit Test!

$$\chi^2 = \frac{(75-90)^2}{75} + \frac{(25-10)^2}{25} + \frac{(75-60)^2}{75} + \frac{(25-40)^2}{25} = 24$$

What is df in this case?

$df = (\# \text{of rows} - 1)(\# \text{colums} - 1) = (2-1)(2-1) = 1\cdot 1 = 1.$

Then $C_{95} = 3.841$. Since $\chi^2 = 24 > 3.841 = C_{95}$, we reject!

This means the variables are not independent.

Multiple Comparison

If you run multiple hypothesis tests, the odds of error become higher.

Lets say:

A_1 = No type 1 error (T1E) in test 1

A_2 = No T1E in test 2

Testing at $\alpha = 0.05$ (95% confidence)

	Probabity of T1E
$P(A_1) = 0.95$	$1 - P(A_1) = 0.05$
$P(A_1 \text{ and } A_2) = P(A_1) \cdot P(A_2)$	$1 - P(A_1 \text{ and } A_2) = 1 - (0.95 \cdot 0.95)$
$= 0.95 \cdot 0.95$	$= 1 - 0.91$
$= 0.91$	$= 0.09$ or $1 - (0.95)^2$

Let m be the number of tests you run.

confidence	$1-\alpha$	$P(A_1)$	} single test
error	$1-(1-\alpha)$	$1-P(A_1)$	
Family wise error rate (FWER)	$1-(1-\alpha)^2$	$1-P(A_1)P(A_2)$	← two tests
	$\vdots$	$\vdots$	
	$1-(1-\alpha)^m$	$1-[P(A_m)^m]$	← m tests

Family wise error rate: The odds you committed a type 1 error in at least one of your tests

How to fix this? Bonferoni correction: For $\alpha = 0.05$, using $\alpha^* = \alpha/m$

$$\Rightarrow FWER = 1 - (1 - \alpha/m)^m \approx \alpha$$