Reminder: The χ^2 Goodness of fit test allows us to use variables from almost any Distribution. So far, we have used the Uniform distribution (for the births of hockey players example and even Mendel's data (which is not a distribution but we do have expected counts under the null hypothesis)

In the next two videos, we will see how we can use The χ^2 Goodness of fit test on **binomial distributed** data (see? The Binomial Distribution is not just used when using the binomial test) and, in the next video, we will see the important **Poisson Distribution**, a distribution that is usually considered to model rare events (like volcanoes erupting or mutations appearing in any particular stretch of a genome)

Example: Is the distribution of gender what would be expected under a random model?

- Gender per birth is independent and equal

Number of boys	Obs num (2 child) families
0	530
1	1332
2	582
Total	2444

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

- Gender per birth is independent and equal

Number of boys	**Obs num (2 child) families**
0	530
1	1332
2	582
Total	2444

k = # of successes	0	1	2
P(*X*=*k*)	0.25	0.50	0.25

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

First we need to estimate $\hat{p}$

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

First we need to estimate $\hat{p}$

$\hat{p}$ = # male children/ all children

= (1332+2*582)/(2444*2)

= 2496/4888 = 0.51

Then, we use this to calculate the expected values of the categories:

$$P[1boy] = \binom{2}{1}(0.5106)^1(1-0.5106)^1 = 0.499$$

χ^2 Goodness of Fit test

<u>Example:</u> Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

First we need to estimate $\hat{p}$

$\hat{p}$ = 2496/4888 = 0.51

Then, we use this to calculate the expected values of the three categories (here is the full calculation for the 1 boy category):

$$P[1boy] = \binom{2}{1}(0.5106)^1(1-0.5106)^1 = 0.499$$

Expected (1 boy families) = 2444x0.499 = 1221.4

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

Num boys	Obs num fam	Exp num fam
0	530	585.3
1	1332	1221.4
2	585	637.3
Total	2444	2444.0

Shortage of two-child families with 0 or 2 and an excess of families with 1. Can this be explained by chance?

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

Test statistic:

$$\chi^2 = \frac{(530-585.3)^2}{585.3} + \frac{(1332-1221.4)^2}{1221.4} + \frac{(582-637.3)^2}{637.3} = 20.04$$

df = # categories - 1- parameter estimated

= 3 - 1 - 1 = 1

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

Test statistic:

$$\chi^2 = 20.04$$

Critical Value:

$$\chi^2_{0.05,1} = 3.84$$

χ^2 Goodness of Fit test

Example: Is the distribution of gender what would be expected under a random model?

H_0: The number of boys in two child families has a binomial distribution

H_A: The number of boys in two child families **does not** have a binomial distribution

Test statistic:

$$\chi^2 = 20.04$$

Critical Value:

$$\chi^2_{0.05,1} = 3.84$$

Conclusion:

Since $P < 0.05$, we reject H_0. The frequency distribution of boys and girls in two-child families does not follow the binomial distribution.