"I am the very model of a modern Major-General;

I've information vegetable, animal and mineral;

I know the Kings of England, and I quote the fights historical,
From Marathon to Waterloo, in order categorical."

William S. Gilbert, English Lyricist, 1836-1911

COHEN CHAP 19 & 20. CATEGORICAL

For EDUC/PSY 6600

MOTIVATING EXAMPLES

Dr. Fisel wishes to know whether a random sample of adolescents will prefer a new of formulation of 'JUMP' softdrink over the old formulation. The ***proportion*** *choosing the new formulation is tested against a hypothesized value of 50%.*

Dr. Sheary hypothesizes that 1/3 of women experience increased depressive symptoms following childbirth, 1/3 experience increases in elevated mood after childbirth, and 1/3 experience no change. To evaluate this hypothesis Dr. Sheary randomly samples 100 women visiting a prenatal clinic and asks them to complete the Beck Depression Inventory. She then re-administers the BDI to each mother one week following the birth of her child. Each mother is classified into one of the 3 previously mentioned categories and ***observed proportions*** *are compared to the* ***hypothesized proportions****.*

Dr. Evanson asks a random sample of individuals whether they see both a physician and a dentist regularly (at least once per year). He compares the ***distributions of these binary variables*** *to determine whether there is a relationship.*

CATEGORICAL METHODS

Instead of means, comparing **counts** and **proportions** within and across groups

- E.g., # ill across different treatment groups

Associations / dependencies among categorical variables

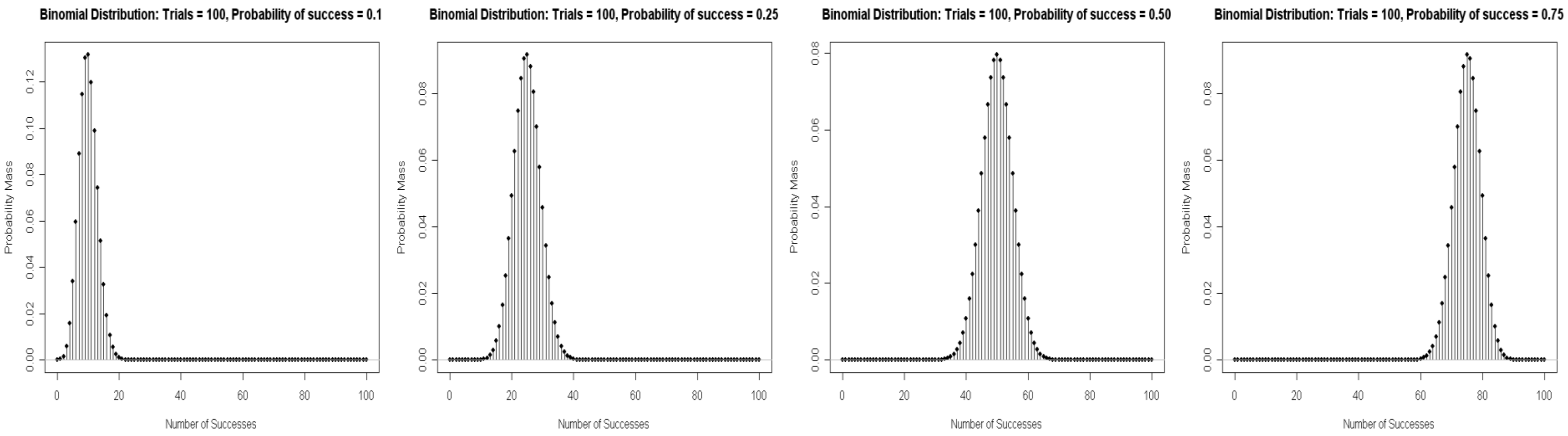
Data are **nominal** or **ordinal**

Discrete probability distribution

- Number of finite values as opposed to infinite

Each subject/event assumes 1 of 2 **mutually exclusive** values (binary or dichotomous)

- Yes/No
- Male/Female
- Well/Ill



THE BINOMIAL DISTRIBUTION: EQ & COIN EXAMPLE

$$p(X) = \frac{N!}{X!(N-X)!} P^X Q^{(N-X)}$$

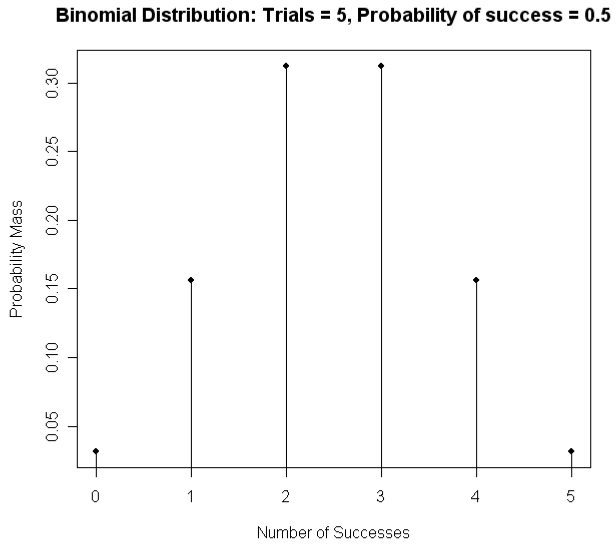
- N = # events
- X = # "successes"
- $P = p$("success")
  - Hypothesized proportion / probability of success
- $Q = p$("failure")
  - Hypothesized proportion / probability of failure
- $P + Q = 1$

- Remember: 0! = 1; $x^0 = 1$

- (Arbitrarily) assign 1 outcome as 'success' and other as 'failure'
- <u>Example</u>: Probability of correctly guessing side of coin 4 out of 5 flips?
  - 5 events, 4 successes, 1 failure
  - $P = p$(correct guess on each flip) = .50
  - $Q = p$(incorrect guess on each flip) = .50

<u>Use equation to obtain:</u>
5 out of 5 successes = .03
4 out of 5 successes = .16
3 out of 5 successes = .31
2 out of 5 successes = .31
1 out of 5 successes = .16
0 out of 5 successes = .03

Sum of probabilities = 1.0



SAMPLING DISTRIBUTION FOR THE BINOMIAL DIST

- Binomial probability distribution for $N = 5$ events, and $P = .5$
- Binomial Distribution Table (exact values)
- Sampling distribution as it was derived mathematically
  - We can only reject H_0 with 0 or 5 out of 5 successes (1-tailed)

Sampling Distribution

$$mean = NP$$
$$variance = NPQ$$
$$SD = \sqrt{NPQ}$$
$$SE_{MEAN} = \sqrt{\frac{PQ}{N}}$$

Example

$M = 5*.5 = 2.5$ (See Histogram)

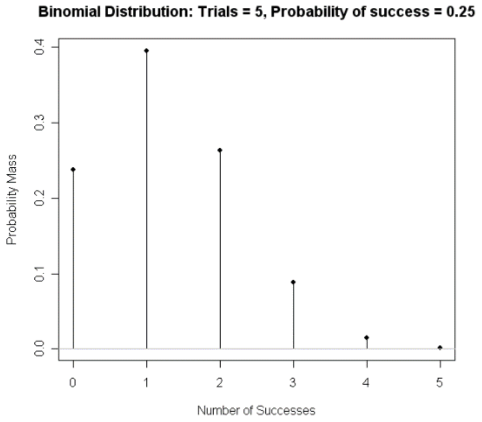
$VAR = 5*.5*.5 = 1.25$

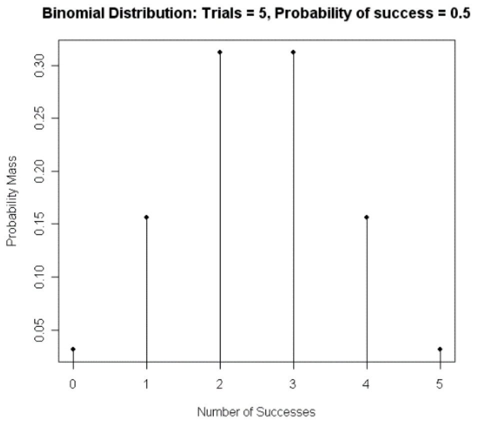
$SD =$ sqrt(1.25) $= 1.12$

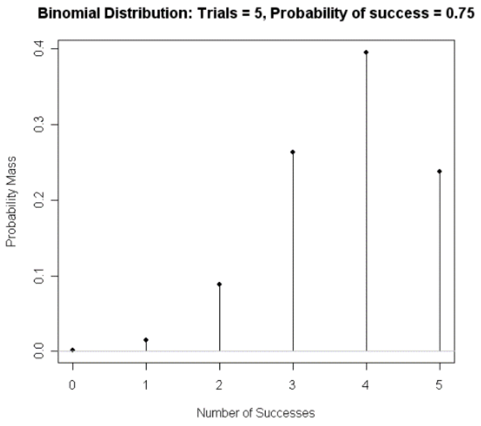
Different binomial distribution for each N

Normal when $P = .50$, skewed when $P \neq .50$

Critical value depends on: N events, X successes, P







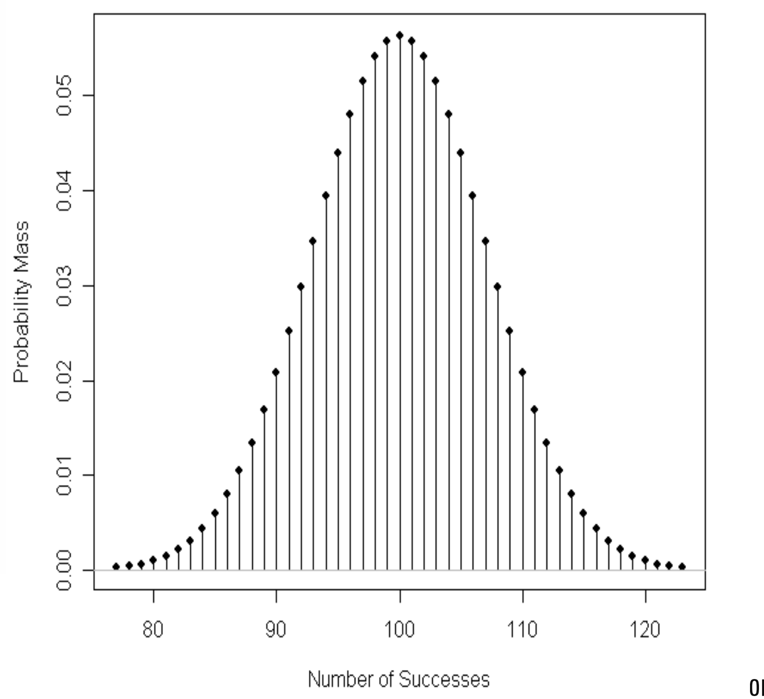
AS N INCREASES, BINOMIAL DISTRIBUTION → NORMAL

n	X	p	n	X	p	n	X	p
1	0	.5000		1	.0176	13	0	.0001
	1	.5000		2	.0703		1	.0016
2	0	.2500		3	.1641		2	.0095
	1	.5000		4	.2461		3	.0349
	2	.2500		5	.2461		4	.0873
3	0	.1250		6	.1641		5	.1571
	1	.3750		7	.0703		6	.2095
	2	.3750		8	.0176		7	.2095
	3	.1250		9	.0020		8	.1571
4	0	.0625	10	0	.0010		9	.0873
	1	.2500		1	.0098		10	.0349
	2	.3750		2	.0439		11	.0095
	3	.2500		3	.1172		12	.0016
	4	.0625		4	.2051		13	.0001
5	0	.0312		5	.2461	14	0	.0001
	1	.1562		6	.2051		1	.0009
	2	.3125		7	.1172		2	.0056
	3	.3125		8	.0439		3	.0222
	4	.1562		9	.0098		4	.0611
	5	.0312		10	.0010		5	.1222

Table A.13
Probabilities of the Binomial Distribution for $P = .5$

"Equally Likely"
Means p = 0.5

Binomial Distribution: Trials = 200, Probability of success = 0.5



BINOMIAL SIGN TEST

Single sample test with binary/dichotomous data

Proportion or % of 'successes' differ from chance?

- H_0: % of observations in one of two categories equals a **specified** % in population
  - H_0: Proportion of 'yes' votes = 50% in population

Assumptions

- Random selection of events or participants
- Mutually exclusive categories
- Probability of each outcome is same for all trials/observations of experiment

- Experiment: Coin flipped 10x, heads 8x
  - Is coin **biased** (Heads > .50)?
- Experiment: 10 women surveyed, 8 select perfume A
  - Is one perfume preferred **over another**?
- For both:
  - H_0: *Proportion* (X) = .50 in population
  - H_1: *Proportion* (X) ≠ .50 in population (2-tailed)

BINOMIAL SIGN TEST: EXAMPLE

- Is occurrence of 8 or more observations in either of the 2 categories unusual?
  - Probability of occurrence given H_0 true in pop.?

n	X	p
	1	.0176
	2	.0703
	3	.1641
	4	.2461
	5	.2461
	6	.1641
	7	.0703
	8	.0176
	9	.0020
10	0	.0010
	1	.0098
	2	.0439
	3	.1172
	4	.2051
	5	.2461
	6	.2051
	7	.1172
	8	.0439
	9	.0098
	10	.0010

NORMAL APPROXIMATION TO THE BINOMIAL I.E."Z-TEST" FOR A SINGLE PROPORTION

What if N were larger, say 15?

- Same proportions: 80% (12/15) Heads & Perfume A
- Sum p(12, 13, 14, 15/15) = .0178 (1-tailed p-value)

Reject H_0 under both 1- and 2-tailed tests

- 2-tailed p = .0178 x 2 = .0356

- Earlier: Binomial distribution → normal distribution, as N → infinity
- Recommendation: Use z-test for single proportion when N is *large* (>25-30)
  - When NP and NQ are both > 10, close to normal
- H_0 and H_1 are same as Binomial Test
- Test statistic:

$$z = \frac{X - PN}{\sqrt{NPQ}} = \frac{p_1 - P}{\sqrt{\frac{PQ}{N}}}$$

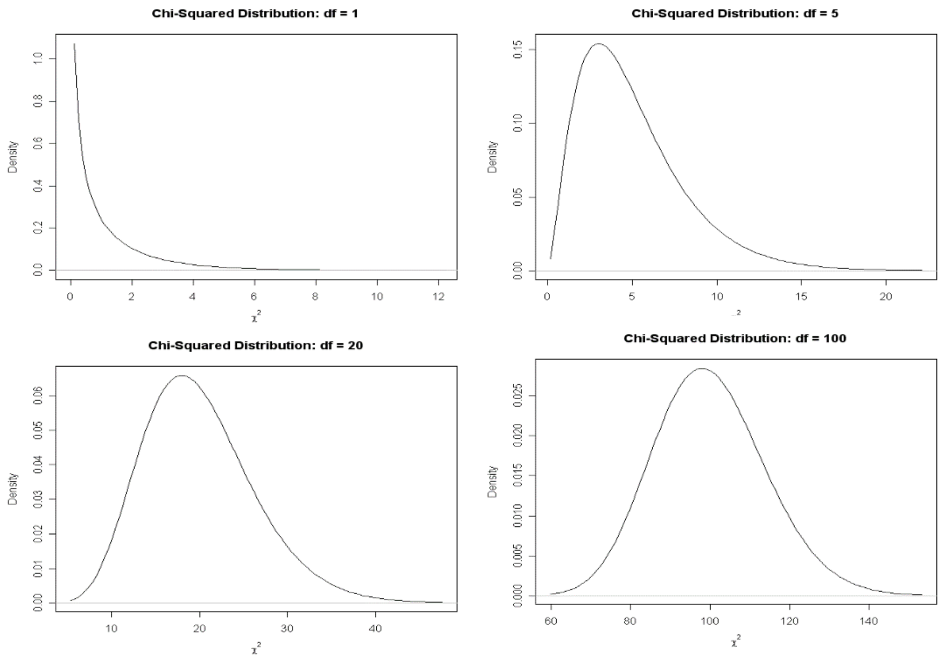
Experiment:
Senator supports bill favoring stem cell research. However, she realizes her vote could influence whether or not her constituents endorse her bid for re-election. She decides to vote for the bill only if 50% of her constituents support this type of research. In a random survey of 200 constituents, 96 are in favor of stem cell research.

Will the senator support the bill?

Standardize:

$$z = \frac{X - NP}{\sqrt{NPQ}}$$

CHI-SQUARE (X^2) DISTRIBUTION



- Family of distributions
  - As df (or k categories) ↑
    - Distribution becomes more normal, bell-shaped
    - Mean & variance ↑
      - Mean = df
      - Variance = 2* df
- $z^2 = \chi^2$
  - Always positive, 0 to infinity
  - 1-tailed distribution
- χ^2 distribution used in many statistical tests

<u>"GOODNESS OF FIT" Testing:</u>

Are <u>observed</u> frequencies **similar** to frequencies <u>expected</u> by chance?

<u>Expected frequencies</u>

Frequencies you'd <u>expect</u> if H_0 were true

Usually equal across categories of variable (N / k)

Can be unequal if theory dictates

CHI-SQUARED: GOODNESS OF FIT TESTS "GOF"

Hypotheses

- H_0: Observed = Expected frequencies in population
- H_1: Observed ≠ Expected frequencies in populatioN

Assumptions

Independent random sample

Mutually exclusive categories

Expected frequencies: ≥ 5 per each cell

General form:

- O = observed frequency
- E = expected frequency

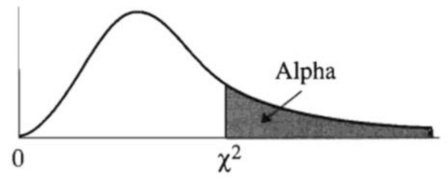
If H_0 were true, numerator would be small

Denominator standardizes difference in terms of expected frequencies

***Aka*: Pearson or '1-way' χ^2 test**

- 1 nominal variable
- 2 or more categories

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$



	ALPHA (AREA IN THE UPPER TAIL)				
df	.10	.05	.025	.01	.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.35	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55
7	12.02	14.07	16.01	18.48	20.28
8	13.36	15.51	17.54	20.09	21.96
9	14.68	16.92	19.02	21.67	23.59
10	15.99	18.31	20.48	23.21	25.19
11	17.28	19.68	21.92	24.72	26.75
12	18.55	21.03	23.34	26.22	28.30
13	19.81	22.36	24.74	27.69	29.82
14	21.06	23.69	26.12	29.14	31.32
15	22.31	25.00	27.49	30.58	32.80
16	23.54	26.30	28.85	32.00	34.27
17	24.77	27.59	30.19	33.41	35.72

If **nominal variable ONLY has 2 categories**, χ^2 GoF test:

- Is another large sample approximation to Binomial Sign Test
- Gives same results as z-test for single proportion as $z^2 = \chi^2$
- Has same H_0 and H_1 as binomial or z-tests

Compare obtained χ^2 statistic to critical value based on df **= k – 1**, k = # categories

GOODNESS OF FIT TESTS – EXAMPLE: K = 2

Hypotheses:
- H_0: P = 0.50
- Observed frequencies: 96 and 104
- Expected frequencies: N / k =200/2 = 100$df = 2 - 1 = 1$

ALWAYS USE COUNTS!!!	1 = "success"	0 = "failure"
OBSERVED (the data)	96	
EXPECTED (based on N, P, Q)		

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

Test Statistic:

$\chi^2_{\ OBSERVED} =$

Critical Value:

χ^2_{CRIT} (__) =

Conclusion:

Note:

GOODNESS OF FIT TESTS – EXAMPLE: K > 2

(ANY NUMBER OF CATEGORIES WITHIN 1 VARIABLE)

ALWAYS USE COUNTS!!!

Hypotheses:

- H_0: " equally likely" (k = 6 & N = 120)
- Expected frequencies: N / k =120/6 = 20
- Observed frequencies: 20, 14, 18, 17, 22, 29 {Mon – Sat}
- $df = 6 - 1 = 5$

	M	T	W	Th	F	S
OBS	20	14	18	17	22	29
EXP						

Test Statistic:

$\chi^2_{\ OBSERVED} =$

Critical Value:

χ^2_{CRIT} (__) =

QUESTION:

Is there a difference in # books checked out for different days of the week?

Conclusion:

We do NOT have evidence the # of books checked out is NOT the same EVERY day

GOODNESS OF FIT TESTS: CONFIDENCE INTERVALS

- ***CI*s for proportions**
  - If $k > 2$, original table converted into table with 2 cells
    - Proportion for category of interest vs proportion in **all other** categories
  - Use same formula for z-test for single proportion:

$$\boxed{P_{obs} \pm z_{crit} \times \sqrt{\frac{P_{obs} \times Q_{obs}}{N}}}$$

- **Say we wanted a *CI* for proportion of books from Saturday (29/120=0.242)**

GOODNESS OF FIT TESTS: EFFECT SIZE

$$\chi^2_{Effect\ Size} = \frac{\chi^2}{N(k-1)}$$

Ranges from 0 to 1

- 0: Expected = Observed frequencies exactly
- 1: Expected ≠ Observed frequencies as much as possible

GOODNESS OF FIT TESTS: POST HOC PAIRWISE TESTS

Like ANOVA, omnibus test, but where do differences lie?

- ‘Pinpointing the action’ in contingency tables
- Post-hoc Binomial, z-tests, or smaller 1-way χ^2 tests
  - Collapsing, ignoring levels
  - Bonferonni correction, more conservative α per comparison
- Examining
  - Observed *vs.* expected frequencies per cell
  - Contributions to χ^2 per cell
- Visual analysis of differences in proportions

2-WAY PEARSON X^2 TEST OF "INDEPENDENCE" OR "ASSOCIATION"

Aka: Contingency table, cross-tabulation, or *row* x *column* ($r \times c$) analysis

- > 1 nominal <u>variable</u>

Is distribution of 1 variable *contingent* on distribution of another?

- Is there an association or dependence between 2 categorical variables

Extension of χ^2 Goodness of Fit Test

<u>Hypotheses:</u>

- H_0: Variables are independent in population
- H_1: Variables are dependent in population

Again, χ^2_{obt} is compared with χ^2_{crit} → $df = (r\text{-}1)(c\text{-}1)$

2-WAY PEARSON X^2 TEST OF "INDEPENDENCE" OR "ASSOCIATION"

Same equation: Standardized squared deviations summed for all cells

$$\chi^2 = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Different method for computing E

- For each cell: Multiply corresponding row and column totals (marginals), divide by N
  -

$$E_{Cell_A} = \frac{(a+b)(a+c)}{N}$$

	Var1		
Var2	**a**	**b**	**a + b**
	c	**d**	**c + d**
	a + c	**b + d**	**a + b + c + d =** N

$$\boxed{EXP_{cell} = \frac{Total_{row} \times Total_{column}}{Total_{grand}}}$$

X^2 TEST OF "INDEPENDENCE" – EXAMPLE:

Experiment:

Random sample of 200 inmates are surveyed about abuse and violent criminal histories

- Relationship between history of abuse and violent crime?

H_0: **No association** between abuse history and violent criminal history in population of prison inmates

- $O_{ij} = E_{ij}$ for all cells in population

H_1: **Association** between abuse history and violent criminal history in population of prison inmates

- $O_{ij} \neq E_{ij}$ for at least one cell in population

Observed frequencies

	Violent Crime		
Abuse	**Yes**	**No**	**Row Sum**
Yes	70	30	100
No	40	60	100
Column Sum	110	90	200

Expected frequencies:

Test Statistic:

APA format: