

Biostatistics

Black pen = Powerpoint

Slide title

Red pen = Equations

Lecture 1 1/14/2014

Some notes on the class:

- Will have colored handouts almost every class
- Most handouts are not posted to Sakai so go to class

START MATERIAL FOR EXAM 1

Introduction to Biostatistics

• statistics = scientific study of numerical data

• biostatistics = application of statistical methods for solution of biological problems

• datum = single measurement

• data = plural measurements

- The first wave of statistical analysis occurred in 1700s and involved taxes and gambling
- Renaissance of biological statistics in 1850s due to work of Darwin, Mendel, Wallace, Sir Francis Galton (Father of Biostatistics)

Basic Statistics

• population = all individuals in a group under study, defined by the researcher

• sample = subset of the population used to estimate properties of the whole population

• composed of individual observations or items

• sampling maximizes information gained per unit effort expended

• variable = property by which individuals differ in some quantifiable way

• continuous variables assume any value between two fixed points

• discrete variables are fixed numeric values

1/16/2014 Population and Samples

SIDENOTE:

Statistics helps:

1. discipline scientific thinking → better scientific conclusions
2. design better experiments/studies

• what 2 properties of a population is a sample used to estimate?

• Mean = average value

• Variance = measure of how variable a population is

• parameter = feature of population

• estimator = feature of sample

	Mean	Variance	Standard Deviation	Size
Parameter	μ	σ^2	σ	N
Estimator	\bar{X}	s^2	s	n

1/16/2014

- As $n \rightarrow N$, $\bar{x} \rightarrow \mu$ and $s^2 \rightarrow \sigma^2$

↳ you get better estimation of population parameter as sample n approaches N

- Statistical inference = process of drawing conclusions about the population from information obtained from the sample

Proper Sampling

• proper sampling = requires that every member of pop. has equal, independent chance of being included in sample

randomization - process of taking random sample

What are accuracy and precision?

• accuracy = nearness of a measurement to actual value of variable being measured

↳ ex: closeness to actual mean, median, etc.

↳ sig figs denote accuracy of measurement

• precision = closeness to each other of repeat measurements of same quantity

↳ A lot of this depends on the measuring instrument

What parameters do we try to measure with samples?

• LOCATION or the central tendency of the data

• DISPERSION or scatter about the central tendency

↳ often displayed as frequency distribution

Measures of Location

• mean = sum of individual observations divided by # of observations

Sample: $\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

Population: $\mu = \frac{\sum_{i=1}^N x_i}{N}$

Same eq, different symbols

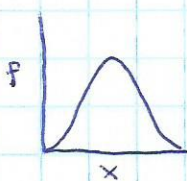
• median = value of variable that has an equal number of items on either side of it

for odd # items: $Md = \left(\frac{n+1}{2}\right)^{\text{th}}$ observation is Median

for even # items: $Md =$ average of middle two items

↳ median might be better in case where there are outliers

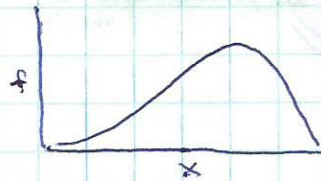
• mode = value that occurs most often



unimodal



bimodal



1/14/14

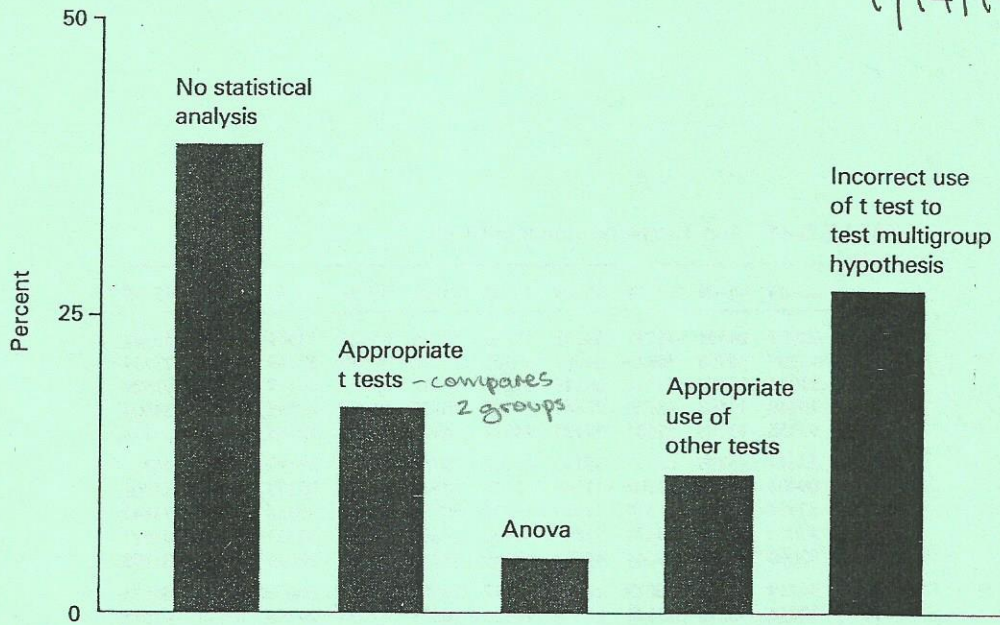


Figure 4-1 Of 142 original articles published in Vol. 56 of *Circulation* (excluding radiology, clinicopathologic, and case reports), 39 percent did not use statistics; 34 percent used a *t* test appropriately to compare two groups, analysis of variance (ANOVA), or other methods; and 27 percent used the *t* test incorrectly to compare more than two groups with each other. (From S. A. Glantz, "How to Detect, Correct, and Prevent Errors in the Medical Literature," *Circulation*, 61:1-7, 1980. By permission of the American Heart Association, Inc.)

Table 1. Statistical Methods Used in 239 Original Research Articles in 6 General Medical Journals, 2005

Type of Test	No. (%)
Descriptive statistics ^a	219 (91.6)
Simple statistics	120 (50.2)
χ^2 Analysis	70 (29.3)
<i>t</i> Test	48 (20.1)
Kaplan-Meier analysis	48 (20.1)
Wilcoxon rank sum test	38 (15.9)
Fisher exact test	33 (13.8)
Analysis of variance	21 (8.8)
Correlation	16 (6.7)
Multivariate statistics	164 (68.6)
Cox proportional hazards	64 (26.8)
Multiple logistic regression	54 (22.6)
Multiple linear regression	7 (2.9)
Other regression analyses ^b	38 (15.9)
None	5 (2.1)
Other methods, techniques, or strategies	
Intention-to-treat analysis	42 (17.6)
Incidence/prevalence	39 (16.3)
Relative risk/risk ratio	29 (12.2)
Sensitivity analyses	21 (8.8)
Sensitivity/specificity	15 (6.3)

^aDescriptive statistics included mean, median, frequency, standard deviation, and interquartile range.

^bOther regression analyses included weighted logistic regression, unconditional logistic regression, conditional logistic regression, longitudinal regression, Poisson regression, pooled logistic regression, nonlinear regression, meta-regression, negative binomial regression, and generalized estimating equations.

TABLE B.41 Ten Thousand Random Digits

	00-04	05-09	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
00	22808	04391	45529	53968	57136	98228	85485	13801	68194	56382
01	49305	36965	44849	64987	59501	35141	50159	57369	76913	75739
02	81934	19920	73316	69243	69605	17022	53264	83417	55193	92929
03	10840	13508	48120	22467	54505	70536	91206	81038	22418	34800
04	99555	73289	59605	37105	24621	44100	72832	12268	97089	68112
05	32677	45709	62337	35132	45128	96761	08745	53388	98353	46724
06	09401	75407	27704	11569	52842	83543	44750	03177	50511	15301
07	73424	31711	65519	74869	56744	40864	75315	89866	96563	75142
08	37075	81378	59472	71858	86903	66860	03757	32723	54273	45477
09	02060	37158	55244	44812	45369	78939	08048	28036	40946	03898
10	94719	43565	40028	79866	43137	28063	52513	66405	71511	66135
11	70234	48272	59621	88778	16536	36505	41724	24776	63971	01685
12	07972	71752	92745	86465	01845	27416	50519	48458	68460	63113
13	58521	64882	26993	48104	61307	73933	17214	44827	88306	78177
14	32580	45202	21148	09684	39411	04892	02055	75276	51831	85686
15	88796	30829	35009	22695	23694	11220	71006	26720	39476	60538
16	31525	82746	78935	82980	61236	28940	96341	13790	66247	33839
17	02747	35989	70387	89571	34570	17002	79223	96817	31681	15207
18	46651	28987	20625	61347	63981	41085	67412	29053	00724	14841
19	43598	14436	33521	55637	39789	26560	66404	71802	18763	80560
20	30596	92319	11474	64546	60030	73795	60809	24016	29166	36059
21	56198	64370	85771	62633	78240	05766	32419	35769	14057	80674
22	68266	67544	06464	84956	18431	04015	89049	15098	12018	89338
23	31107	28597	65102	75599	17496	87590	68848	33021	69855	54015
24	37555	05069	38680	87274	55152	21792	77219	48732	03377	01160
25	90463	27249	43845	94391	12145	36882	48906	52336	00780	74407
26	99189	88731	93531	52638	54989	04237	32978	59902	05463	09245
27	37631	74016	89072	59598	55356	27346	80856	80875	52850	36548
28	73829	21651	50141	76142	72303	06694	61697	76662	23745	96282
29	15634	89428	47090	12094	42134	62381	87236	90118	53463	46969
30	00571	45172	78532	63863	98597	15742	41967	11821	91389	07476
31	83374	10184	56384	27050	77700	13875	96607	76479	80535	17454
32	78666	85645	13181	08700	08289	62956	64439	39150	95690	18555
33	47890	88197	21368	65254	35917	54035	83028	84636	38186	50581
34	56238	13559	79344	83198	94642	35165	40188	21456	67024	62771
35	36369	32234	38129	59963	99237	72648	66504	99065	61161	16186
36	42934	34578	28968	74028	42164	56647	76806	61023	33099	48293
37	09010	15226	43474	30174	26727	39317	48508	55438	85336	40762
38	83897	90073	72941	85613	85569	24183	08247	15946	02957	68504
39	82206	01230	93252	89045	25141	91943	75531	87420	99012	80751
40	14175	32992	49046	41272	94040	44929	98531	27712	05106	35242
41	58968	88367	70927	74765	18635	85122	27722	95388	61523	91745
42	62601	04595	76926	11007	67631	64641	07994	04639	39314	83126
43	97030	71165	47032	85021	65554	66774	21560	04121	57297	85415
44	89074	31587	21360	41673	71192	85795	82757	52928	62586	02179
45	07806	81312	81215	99858	26762	28993	74951	64680	50934	32011
46	91540	86466	13229	76624	44092	96604	08590	89705	03424	48033
47	99279	27334	33804	77988	93592	90708	56780	70097	39907	51006
48	63224	05074	83941	25034	43516	22840	35230	66048	80754	46302
49	98361	97513	27529	66419	35328	19738	82366	38573	50967	72754

1/16/2014

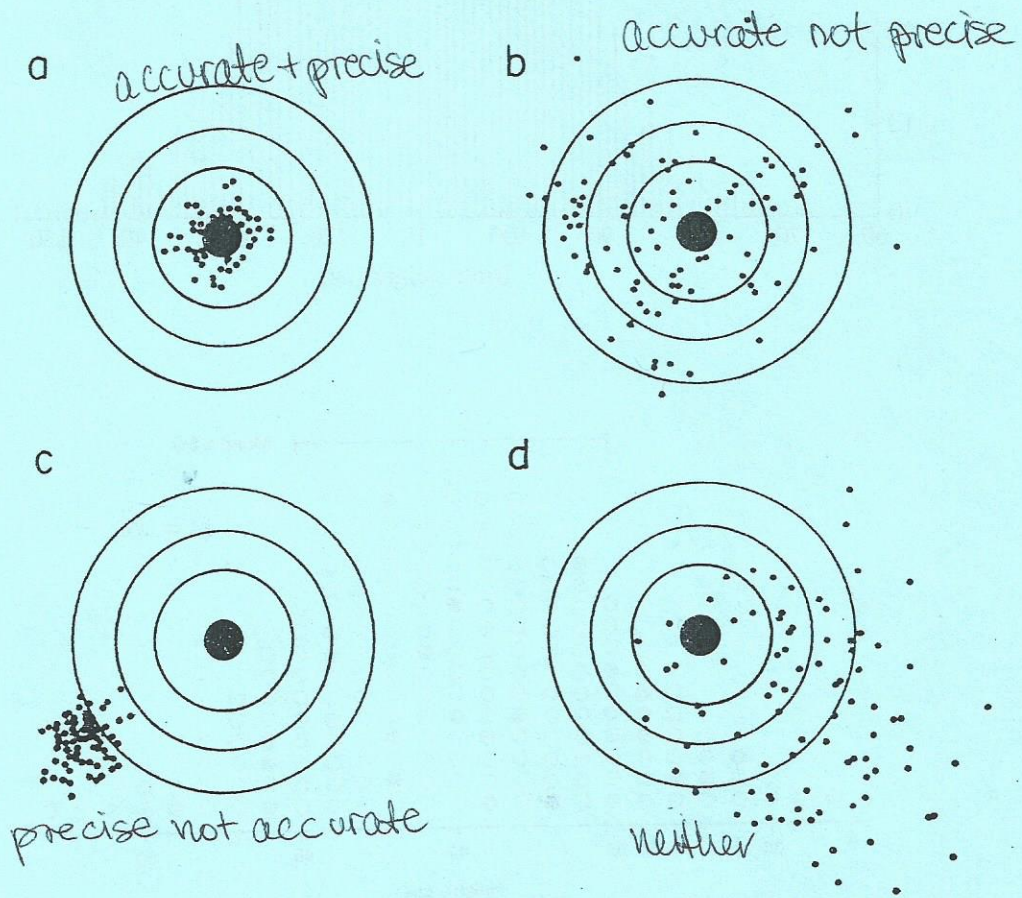


Fig. I.1. Bias (accuracy) and precision target: (a) unbiased and precise, (b) unbiased but not precise, (c) biased but precise, and (d) biased and not precise

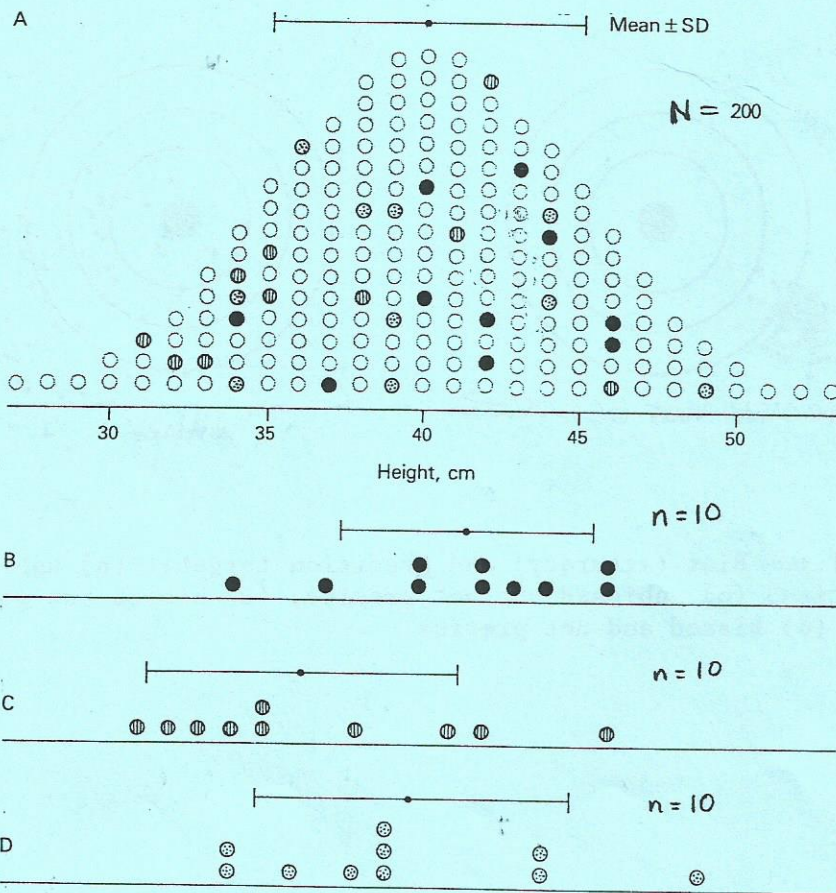
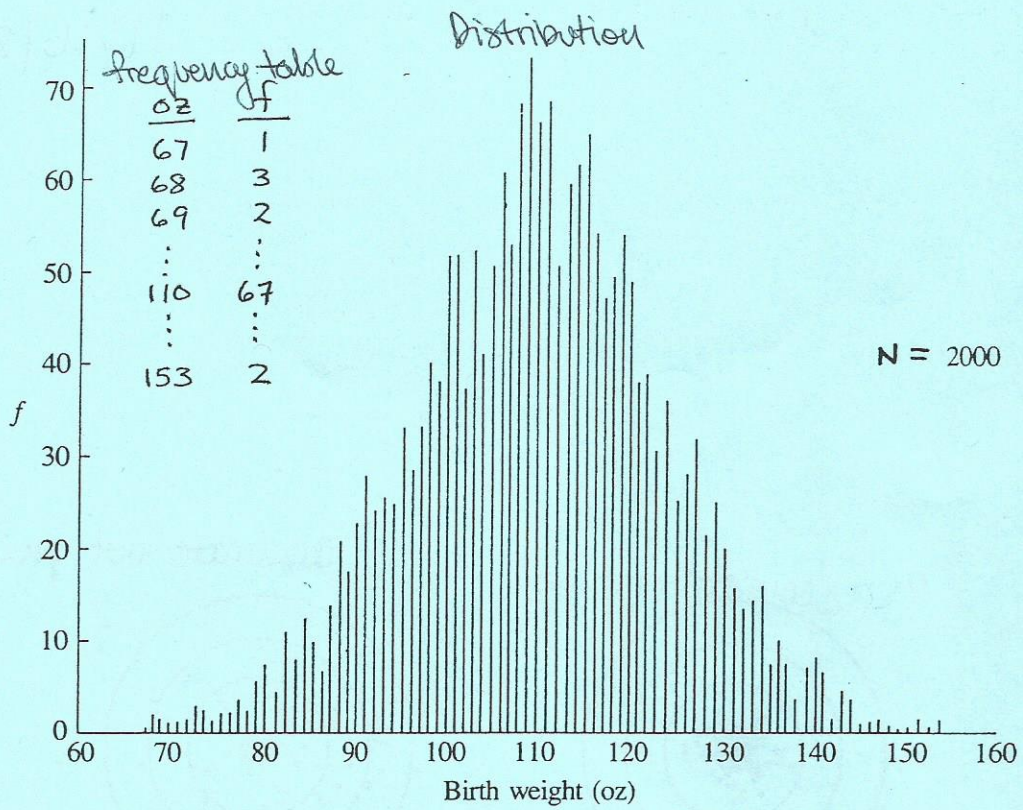
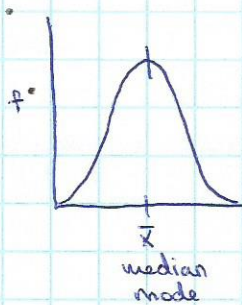


Figure 2-6 If one draws three different samples of 10 members each from a single population, one will obtain three different estimates of the population mean and standard deviation.

1/16/14



parameters all same

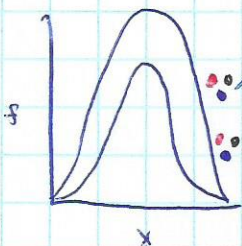
but usually...



all over the place

1/21/2014

Measures of Dispersion



- same mean, different dispersion → need ways to quantify dispersion

range = largest value - smallest value (w)

variance = sum of squared deviations of each observation from the mean divided by the sample size $n-1$

Sample Variance:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Sum of Squares (SS)
freedom (v)

Population Variance:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad \text{RARELY KNOW THIS!}$$

Why do we square the variance?

↳ If we didn't the summation would = 0!

$$\star \sum (x_i - \bar{x}) = (\sum x_i) - (\sum \bar{x}) = (\sum x_i) - (n\bar{x}) = (\sum x_i) - n \left(\frac{\sum x_i}{n} \right) = (\sum x_i) - (\sum x_i) = 0$$

standard deviation = square root of the variance; typical departure from the mean

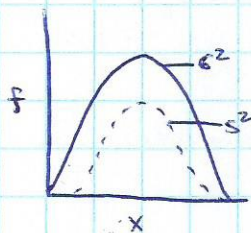
$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

What are degrees of freedom?

↳ you lose 1 degree of freedom for every parameter estimated

↳ Ex. you use \bar{x} to estimate μ so the SD and Variance use $n-1$

- Sample variance tends to underestimate population variance so compensate by dividing $n-1$



- Sample may show correct mean but dispersion may be off

1/21/2014

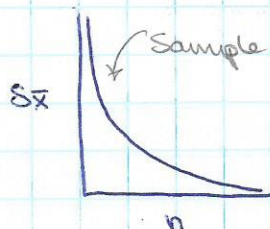
Some Sample Properties

Standard error = "certainty" that a mean computed from random sample estimates true mean of population

* $S_{\bar{x}} = \frac{S}{\sqrt{n}}$ $S_{\bar{x}} \rightarrow 0$ as sample size n increases ($n \rightarrow \infty$ $n \rightarrow N$)

↳ In general, a good sample:

$S_{\bar{x}} \leq 0.1 \bar{x}$ (10% of actual mean)



Sample sizes impacted by $S_{\bar{x}}$ drastically

- Good for $S_{\bar{x}}$ to be low; infinitely approaches 0

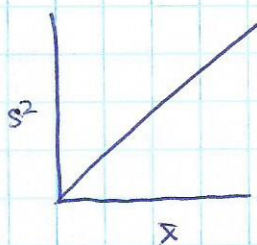
At home problem:

$X_i = 1, 3, 4, 6, 10, 11, 100, 105$

Calculate \bar{x} , median, SD, var

$\bar{x} = 30$ median = 8. var = 2015 SD = 44.9 $S_{\bar{x}} = 16$

Coefficient of variation = metric that allows you to compare variation widely differing biological entities (CV)



- In nature as \bar{x} increases, so does s^2

$CV = \frac{S}{\bar{x}}$ as fraction

$CV = \frac{S}{\bar{x}} \times 100$ (as percent)

↳ Use to compare things by percentages which allows comparison of two sets of unlike data, means

- Example



Comparing variance of tail length b/w mice, elephants

mouse $CV = \frac{S}{\bar{x}} = \frac{2cm}{10cm} \rightarrow 20\%$ elephant = $\frac{0.4m}{2m} \rightarrow 20\%$

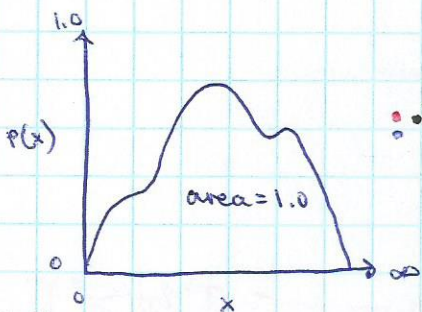
↳ Different organism, same tail length variance

1/21/2014

Variables and Distributions

• **Random variable** = one that is subject to random variation from a variety of sources; i.e. biological variables

↳ can be described by **probability distribution** which is frequency distribution about which predictions can be made



Probability Distributions

• **Discrete distributions** = random variables that take on only a finite and specified # of values or categories

Ex. Binomial Distribution

• For binomials, probability of x, P(x):

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

n = number of trials

x = # of trials w/ specific outcomes

p = outcome probability (Ex. heads)

q = alternative outcome probability

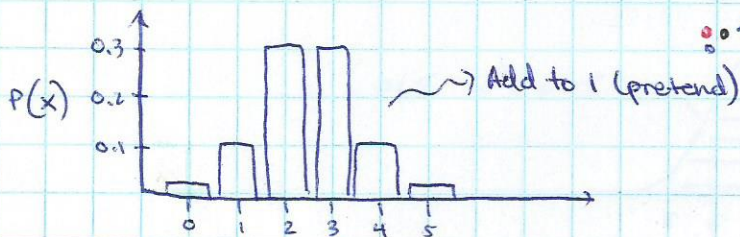
$$\rightarrow q = 1 - p$$

↳ tails

• **Independence** = outcome of one trial or event has no effect on outcome of another trial

Example → Flip fair coin 5 times

n = 5 p = 0.5 (heads) q = 0.5 (tails) x = possible outcomes = 0, 1, 2, 3, 4, 5 heads = 6 outcomes



• **Unique to binomial:**

$$\mu = n \cdot p = 2.5$$

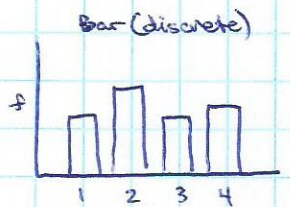
$$\sigma^2 = n \cdot p \cdot q = 1.25$$

How many combinations?

$${}_n C_x = {}_5 C_2 = 5 \text{ trials of 2 possible outcomes} = \frac{n!}{x!(n-x)!} = \frac{5 \times 4}{2 \times 1} = 10$$

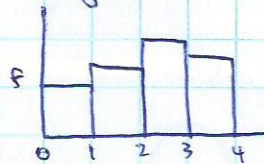
Side Note

- Plotting discrete data, continuous data



bars don't touch!

Histogram (continuous)



bars touch b/c the data are continuous

1/21/2014

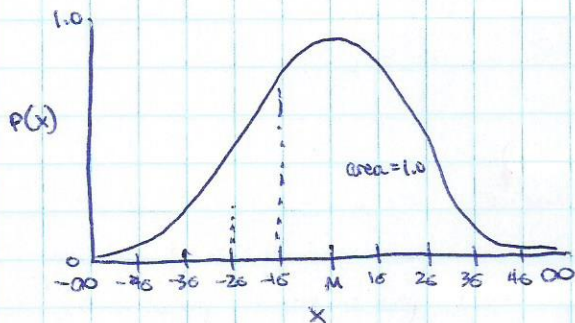
Continuous distributions = apply to random variables that can take on an ∞ number

of values

Ex. Normal distribution, "bell shaped curve"

- 1733 : Abraham de Moivre
- 1774 : Pierre Simon
- 1809 : Karl Gauss
- 1877 : Sir Francis Galton \rightarrow "normal" distribution
- 1920 : Karl Pearson \rightarrow "normal"

Normal (Z-distribution)

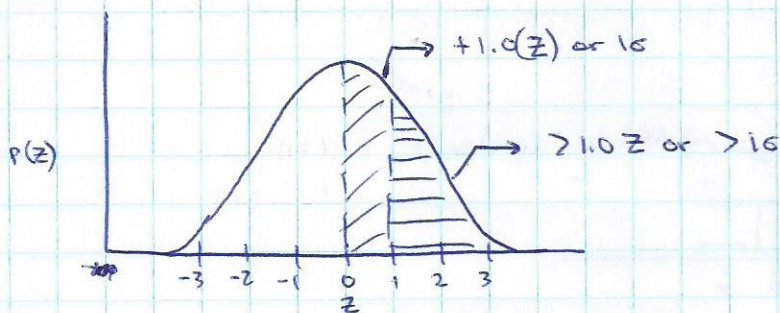


$N = \infty$

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$$

need to know μ and σ to calculate probability associated with any value of x

$$Z = \frac{\text{Samples } (X_i - \mu)}{\sigma} = \frac{\text{Means } (\bar{x} - \mu)}{\sigma_{\bar{x}}} \quad \text{Z-score or standard Deviate}$$



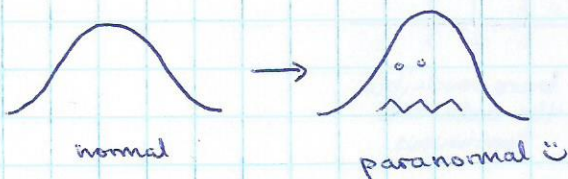
\hookrightarrow you can use the BZ table to find Z scores

Coffee...

1/23/2014

Properties of the Normal Distribution

- extends from $-\infty$ to $+\infty$
- curve is symmetrical about mean
- total probability sums to zero
- n is assumed to be infinite (smooth curve)



1/21/14

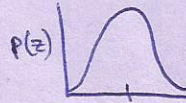


TABLE B.2 Proportions of the Normal Curve (One-Tailed)

This table gives the proportion of the normal curve that lies beyond (i.e., is more extreme than) a given normal deviate; e.g., $Z = (X_i - \mu)/\sigma$ or $Z = (\bar{X} - \mu)/\sigma_{\bar{X}}$. For example, the proportion of a normal distribution for which $Z \geq 1.51$ is 0.0655.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	Z
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641	0.0
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247	0.1
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859	0.2
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483	0.3
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121	0.4
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776	0.5
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451	0.6
0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2207	0.2177	0.2148	0.7
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867	0.8
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611	0.9
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379	1.0
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170	1.1
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985	1.2
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823	1.3
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681	1.4
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559	1.5
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455	1.6
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367	1.7
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294	1.8
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233	1.9
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183	2.0
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143	2.1
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110	2.2
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084	2.3
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064	2.4
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048	2.5
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036	2.6
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026	2.7
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019	2.8
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014	2.9
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010	3.0
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007	3.1
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005	3.2
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	3.3
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	3.4
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	3.5
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.6
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.7
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.8

One-tail 16

26

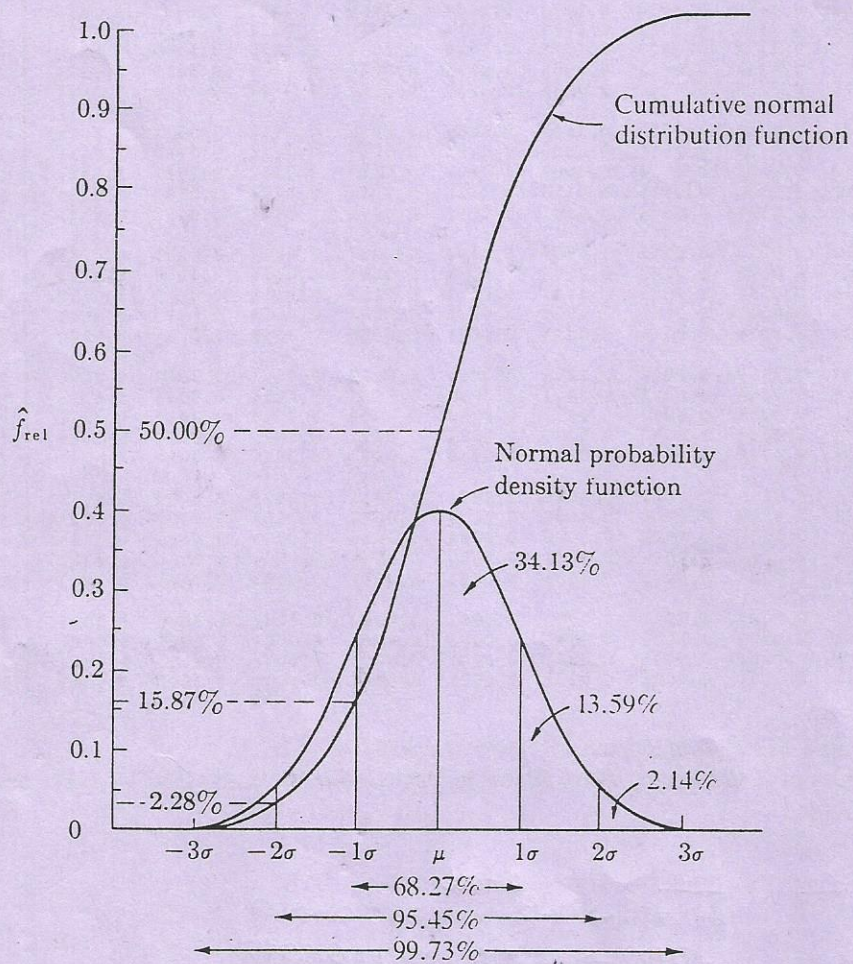
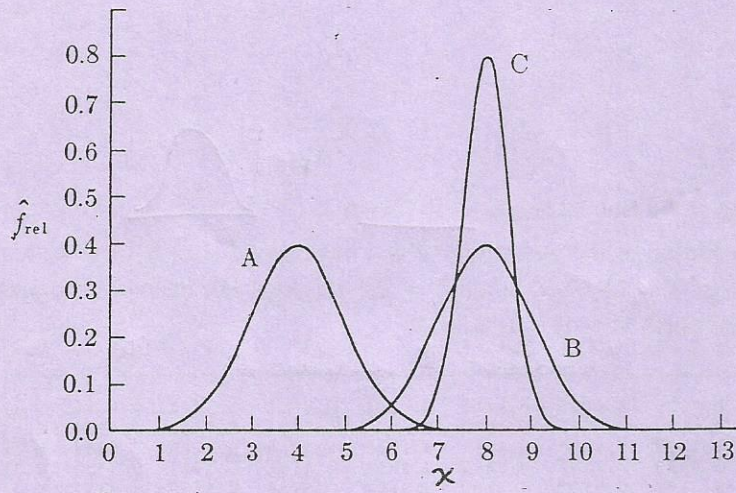
36

Table B.2 was prepared using an algorithm of Hastings (1955: 187). Probabilities for values of Z in between those shown in this table may be obtained by either linear or harmonic interpolation.

At home

<u>Range of X</u>	<u>Area</u>
$\mu \pm 1\sigma$	68.27% $\rightarrow 1 - 2(0.1587) = 0.6827$
$\mu \pm 2\sigma$	95.45% $\rightarrow 1 - 2(0.0228) = 0.9545$
$\mu \pm 3\sigma$	99.73% $\rightarrow 1 - 2(0.0013) = 0.9973$

<u>Area</u>	<u>Range of X</u>
50% \rightarrow 0.5000	$\mu \pm 0$ $0.5 = 1 - 2(x) \rightarrow x = 0.25 \rightarrow$ 0.675 $0.675 \checkmark$
95%	$\mu \pm \sigma$ $0.95 = 1 - 2(x) \rightarrow x = 0.025 \rightarrow$ 1.96 $1.96 \checkmark$
99%	$\mu \pm \sigma$ $0.99 = 1 - 2(x) \rightarrow x = 0.005 \rightarrow$ 2.576 $2.576 \checkmark$



1/23/2014

Example

$N = 1000$ body weights $\mu = 70$ kg $\sigma = 10$ kg

Q: What proportion of pop is > 80 kg

$$Z = \frac{x_i - \mu}{\sigma} = \frac{80 - 70}{10} = 1$$

$$P(x > 80) = P(Z > 1) = 0.1587 \rightarrow \sim 16\%$$

* At home, given above info...

$$P(x) > 85? \quad Z = \frac{85 - 70}{10} = 1.5 \quad P(x > 85) = P(Z > 1.5) = 0.0668 \rightarrow 6.68\% \checkmark$$

Now say $\mu = 70$ kg $\sigma = 5$ kg, $P(x) > 80$ kg?

$$Z = \frac{80 - 70}{5} = 2 \quad P(x > 80) = P(Z > 2) = 0.0228 = 2.3\%$$

$$P(x > 80) = P(Z > 2) = 0.023 \rightarrow 2.3\%$$

* At home $P(x) > 80$ or < 60 — for $\sigma = 10 \rightarrow P(x) > 80 = 16\%$ $P(x) < 60 = 16\% \rightarrow \sim 32\%$

for $\sigma = 5 \rightarrow P(x) > 80 = 2.3\%$ $P(x) < 60 = 2.3\% \rightarrow 4.6\%$

for $\sigma = 10$ $P(x) > 80 = 16\%$ $P(x) > 70 = 50\% + (-1.5 + 1.6) = 34\%$

for $\sigma = 5$ $P(x) > 80 = 2.3\%$ $P(x) > 70 = 50\% \rightarrow 97.7 - 50 = 47.7\%$

for $\sigma = 10$ $P(x) > 80 = 16\%$ $P(x) < 65 = 31\% \rightarrow 1 - (16 + 31) = 53\%$

for $\sigma = 5$ $P(x) > 80 = 2.3\%$ $P(x) < 65 = 16\% \rightarrow 1 - (.16 + .023) = 81.7\%$

- x_i = biological mechanism underlying this variable? \pm Σ errors

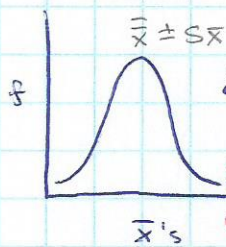
↳ Ex. weight = growth \pm Σ (diet, exercise, etc.)

Central Limit Theorem

holds that the distribution of errors tends to be approximately normal and thus samples drawn from populations are also normal

Recall:

$$s_x = \frac{s}{\sqrt{n}} \text{ (standard error of the mean)}$$



distribution of means will be approximately

normal

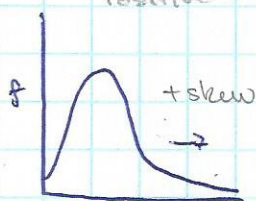
SEM estimates uncertainty of mean

SD estimates variation in individual measurements

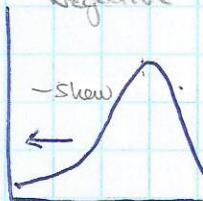
$Z = \frac{\bar{x} - \mu}{s_x}$ The CLT further implies that the distribution of sample means will be approximately normal regardless of original distribution

- Skewness of a population

Positive

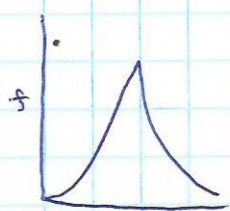


Negative

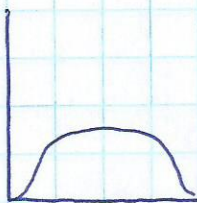


- Where mean is relative to ϕ

$1/2 \cdot 2/14$
 Kurtosis = peakedness of the distribution



- leptokurtosis
 + kurtosis



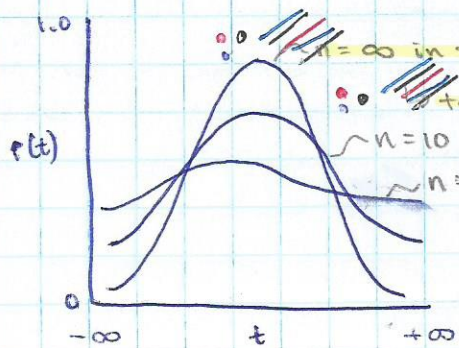
- platykurtosis
 - kurtosis

When do we use the Normal Distribution?

- Applies to population of infinite size
- Can be used for relatively large samples
- But what distribution do we use for small sample sizes?

Student's T-distribution

- Analog of the normal distribution that can be applied to small sample sizes



$n = \infty$ in size \rightarrow approaches $t \rightarrow z$ distribution

tails of t will change w/ smaller n

$n = 10$

$n = 2$

standard deviate for t -dist:

$$t = \frac{\sum (x_i - \bar{x})}{s} \text{ Sample}$$

estimators, rather than params

Important Elements of T-distribution

- for $n = \infty$, t - and z -distributions are identical (very similar for $n > 100$)

- as n declines, tail areas get larger
- t -dist depends on 3 estimators (\bar{x}, s, n)
- recall z -dist depends only on μ, σ

Confidence Limits (of the mean)

$$P(L_1 \leq \mu \leq L_2) = 1 - \alpha \quad 0 \leq \alpha \leq 1 \quad \text{Ex. } \alpha = 0.05$$

↑ lower limit ↑ upper limit

- tells us the probability that μ lies within limits

$$L_1 = \bar{x} - t s_{\bar{x}} \quad L_2 = \bar{x} + t s_{\bar{x}}$$

$$CI = \bar{x} \pm t_{\alpha, \nu} s_{\bar{x}} \quad \text{confidence interval}$$

- this value comes from t -table (pink sheet)
- contains true mean (μ) with $P(x) = 1 - \alpha$

$\nu =$ degrees of freedom

degrees of freedom
2 tail t-dist

1/23/14

TABLE B.3 Critical Values of the *t* Distribution

v	$\alpha(2)$	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
	$\alpha(1)$	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619	
2	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599	
3	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924	
4	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610	
5	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869	
6	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959	
7	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408	
8	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041	
9	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781	
10	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587	
11	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437	
12	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318	
13	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221	
14	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140	
15	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073	
16	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015	
17	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965	
18	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922	
19	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883	
20	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850	
21	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819	
22	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792	
23	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768	
24	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745	
25	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725	
26	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707	
27	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690	
28	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674	
29	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659	
30	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646	
31	0.682	1.309	1.696	2.040	2.453	2.744	3.022	3.375	3.633	
32	0.682	1.309	1.694	2.037	2.449	2.738	3.015	3.365	3.622	
33	0.682	1.308	1.692	2.035	2.445	2.733	3.008	3.356	3.611	
34	0.682	1.307	1.691	2.032	2.441	2.728	3.002	3.348	3.601	
35	0.682	1.306	1.690	2.030	2.438	2.724	2.996	3.340	3.591	
36	0.681	1.306	1.688	2.028	2.434	2.719	2.990	3.333	3.582	
37	0.681	1.305	1.687	2.026	2.431	2.715	2.985	3.326	3.574	
38	0.681	1.304	1.686	2.024	2.429	2.712	2.980	3.319	3.566	
39	0.681	1.304	1.685	2.023	2.426	2.708	2.976	3.313	3.558	
40	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551	
41	0.681	1.303	1.683	2.020	2.421	2.701	2.967	3.301	3.544	
42	0.680	1.302	1.682	2.018	2.418	2.698	2.963	3.296	3.538	
43	0.680	1.302	1.681	2.017	2.416	2.695	2.959	3.291	3.532	
44	0.680	1.301	1.680	2.015	2.414	2.692	2.956	3.286	3.526	
45	0.680	1.301	1.679	2.014	2.412	2.690	2.952	3.281	3.520	
46	0.680	1.300	1.679	2.013	2.410	2.687	2.949	3.277	3.515	
47	0.680	1.300	1.678	2.012	2.408	2.685	2.946	3.273	3.510	
48	0.680	1.299	1.677	2.011	2.407	2.682	2.943	3.269	3.505	
49	0.680	1.299	1.677	2.010	2.405	2.680	2.940	3.265	3.500	
50	0.679	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496	
1000	0.675	1.282	1.646	1.962	2.330	2.581	2.813	3.098	3.300	
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.0902	3.2905	

Basically defines the SD from the mean

This table was prepared using Equations 26.7.3 and 26.7.4 of Zelen and Severo (1964), except for the values at infinity degrees of freedom, which are adapted from White (1970). Except for the values at infinity degrees of freedom, *t* was calculated to eight decimal places and then rounded to three decimal places.

Examples:

$$t_{0.05(2),13} = 2.160 \quad \text{and} \quad t_{0.01(1),19} = 2.539$$

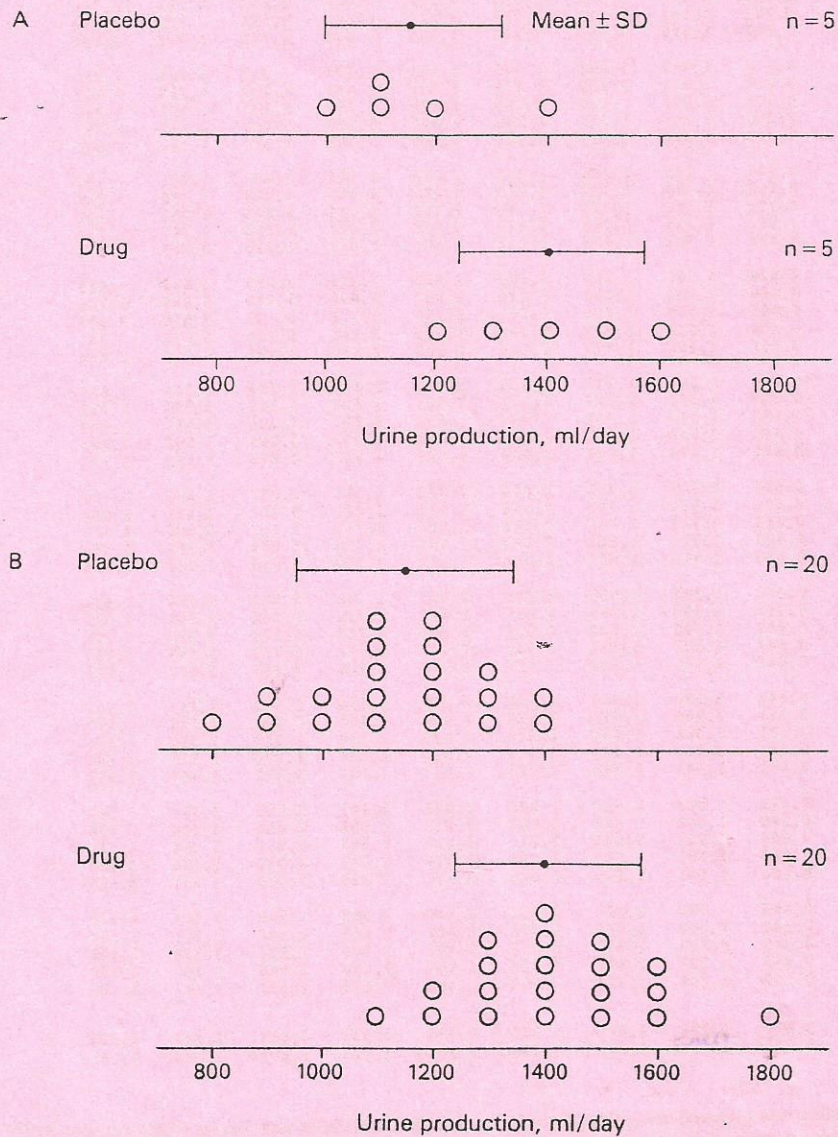
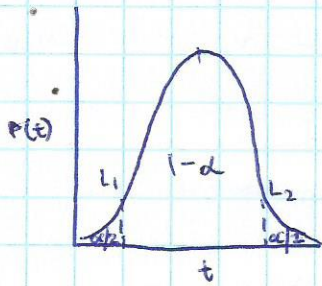


Figure 4-2 (A) Results of a study in which five people were treated with a placebo and five people were treated with a drug thought to increase daily urine production. On the average, the five people who received the drug produced more urine than the placebo group. Are these data convincing evidence that the drug is an effective diuretic? (B) Results of a similar study with 20 people in each treatment group. The means and standard deviations associated with the two groups are similar to the results in panel A. Are these data convincing evidence that the drug is an effective diuretic? If you changed your mind, why did you do it?

1/23/2014



$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} \text{ estimating the mean}$$

$$2(\alpha/2) + (1-\alpha) = 1$$

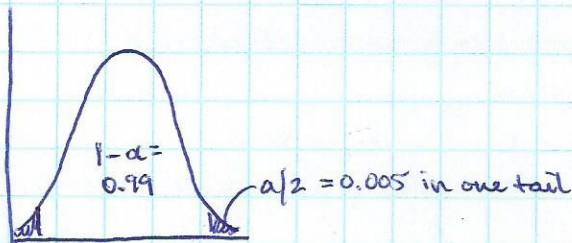
↳ α usually set at 0.05

↳ For $\alpha = 0.05$, with $P(x) = 1 - \alpha$ (95%), the interval L_1 to L_2 will contain the μ , given our n

- called the 95% confidence interval

↳ we're 95% sure this interval contains true population mean

↳ For $\alpha = 0.01$, interval widens to 99% CI



Example

$x_1 =$ days to recovery = 1, 3, 4, 5, 7 days

$$\bar{x} = 4 \quad SD = 2.24 \text{ days} \quad s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{2.24}{\sqrt{5}} = 1$$

↳ What are the 95% confidence limits? $\alpha = 0.05$

Hypothesis: Interval $L_1 - L_2$ contains μ $H_0: L_1 \leq \mu \leq L_2$

- since $n = 5, \nu = 4$

$t_{0.05(2), 4} = 2.776$ * need α & $df(\nu)$ to find this

$$L_1 = \bar{x} - t_{\alpha, \nu} s_{\bar{x}} = 4 - (2.776 \cdot 1) = 1.224$$

$$L_2 = \bar{x} + t_{\alpha, \nu} s_{\bar{x}} = 4 + (2.776 \cdot 1) = 6.776$$

$$P(1.224 \leq \mu \leq 6.776) = 95\%$$

• smaller n allows smaller $s_{\bar{x}}$ and gives more degrees of freedom

• can also constrain the confidence intervals to get smaller range for CI

* At Home

$\uparrow n \rightarrow \downarrow s_{\bar{x}}$ and $\uparrow \nu$ which $\downarrow t$

$n = 10, \mu = 4, s = 2.24$, what is 95% CI? $s_{\bar{x}} = 0.708, t_{0.05(2), 9} = 2.262, L_1 = 4 - 1.60, L_2 = 4 + 1.60 \rightarrow CI = 4 \pm 1.60$

$n = 5, \mu = 4, s = 2.24, \alpha = 0.10$, what is 90% CI? $s_{\bar{x}} = 1.0, t_{0.1(2), 4} = 2.132, CI = 4 \pm 2.132$

↳ all the same but $\alpha = 0.02$, what is 98% CI? $s_{\bar{x}} = 1, t_{0.02(2), 4} = 3.747, CI = 4 \pm 3.747$

↳ t_x exceeds range of data bc t -dist goes $-\infty$ to ∞

↳ include $P(L_1 \leq \mu \leq L_2) = 1 - \alpha$ in answers

1/28/2014

↳ Notes from email since class was cancelled

- Another common use for the t-distribution is to test hypotheses about 2 means

Independent (Groups) t-test

• Often t-test are used to analyze experimental results in which there are 2 groups selected randomly and then subjected to two different treatments (one may be a control)

• We want to know if the means of the two groups are significantly different

- We construct a null (no difference) hypothesis for the parametric means of the two groups

• Null Hypothesis: $H_0: \mu_1 = \mu_2$

• Alternative Hypothesis: $H_A: \mu_1 \neq \mu_2$

- We calculate a test statistic to test H_0

↳ This t-stat is compared to t-table critical value

- We set α -level typically to 0.05

- When H_0 is rejected at specific α , we say that the samples, groups, \bar{x} 's are significantly different at probability $p \leq \alpha$ or at $p = \alpha$ precisely

Paired t-test

• Sometimes we want to know if there is a difference b/w paired observations

• We ask if the mean parametric difference (\bar{D}) b/w logically paired observations is equal to zero

• Null Hypothesis: $H_0: \bar{D} = 0$

• Alternative Hypothesis: $H_A: \bar{D} \neq 0$

- Samples occur in pairs

- If we reject H_0 we can conclude that a significant difference exists in the paired observations

• Does not mean the means are different simply that the mean difference is not zero and some consistent deviation exists b/w paired observations

- Examples of paired t-test include before/after comparisons, two measurements on the same item or similar items

Paired t-test Eq:

$$t = \frac{\bar{d} - 0}{s_{\bar{d}}}$$

\bar{d} = mean of differences between sample sets

$$s_{\bar{d}} = \frac{s}{\sqrt{n}}$$

Question: Was there a significant difference in points that Chris Thomas scored in big ND wins versus big ND losses?

$H_0: \mu_W = \mu_L$ $H_a: \mu_W \neq \mu_L$

L points	W points
14	19
5	10
4	32
24	8
8	15
6	13
13	
$\sum = 87$	$\sum = 97$

Back

$n = 8$ $n = 6$

$\bar{X}_L = 10.875$ $\bar{X}_W = 16.17$

$S_{\bar{X}_L - \bar{X}_W} = 1.544$

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{10.875 - 16.17}{1.544} = -3.427$$

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}} = \sqrt{\frac{66.93}{8} + \frac{66.93}{6}} = 1.544$$

$$S_p^2 = \frac{SS_1 + SS_2}{n_1 + n_2} = \frac{304.875 + 446.25}{12} = 66.93$$

$t_{0.05, 12} = 2.176$

$t < t_{0.05, 12}$ so reject H_0

Yes, there was a significant difference

1/30/2014

SB Tribune 1/21/03

COLLEGE

THOMAS IN BIG GAMES

2002-03 Game	FG	3-pt.	Points	Assists	TO	Outcome
1-18 at Kentucky	4-17	2-8	14 ✓	6	9	L (88-73)
1-6 at Pittsburgh	2-15	1-7	5 ✓	5	4	L (72-55)
12-8 vs. Texas (BB&T)	6-18	4-6	19 W	8	3	W (98-92)
12-7 vs. Maryland (BB&T)	3-14	0-4	10 W	9	3	W (79-67)
12-2 Marquette	12-18	5-8	32 W	10	5	W (92-71)
2001-02						
3-16 vs. Duke (NCAA)	2-14	0-6	4 ✓	7	6	L (84-77)
3-14 vs. NC Char. (NCAA)	3-9	2-6	8 W	7	2	W (82-63)
3-8 vs. Conn. (BET)	7-14	6-9	24 ✓	10	5	L (82-77)
3-7 vs. St. John's (BET)	5-16	3-7	15 W	11	5	W (83-63)
3-2 Providence	4-12	2-6	13 W	3	3	W (76-68)
2-17 Syracuse	3-11	1-5	8 ✓	6	3	L (68-65)
2-9 at Rutgers	0-14	0-9	6 ✓	9	5	L (69-62)
1-21 Georgetown	5-13	0-4	13 ✓	8	3	L (83-73)
1-19 Kentucky	4-15	2-11	13 ✓	4	8	L (72-65)
TOTALS:	60-200	28-96	184	103	64	6-8 record
	30%	29.1%	13.1	7.3	4.5	

Bright lights, big player: Thomas in a clutch role

Notre Dame sophomore point guard Chris Thomas has often been referred to as a "bright-lights guy" by Irish head coach Mike Brey because Thomas likes the challenge of competing when the Irish play on the game's biggest stage.

But Thomas again struggled in Saturday's 15-point loss at Ken- tucky, where he committed a ca- reer-high nine turnovers.

Here is a look back at how Thomas has fared in Notre Dame's "bright-lights" games, contests that were against quality non-con- ference opponents or on national television, dating back to last sea- son's first meeting with Kentucky at the Joyce Centre.

L	14	5	4	24	8	6	13	13
M	19	10	32	8	15	13		

Losses $n = 8$ $\bar{x} = 10.875$
 Wins $n = 6$ $\bar{x} = 16.17$

$$t = \frac{|\bar{x}_L - \bar{x}_M|}{S_{\bar{x}_1 - \bar{x}_2}} \quad S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{SP^2}{n_1} + \frac{SP^2}{n_2}} \quad SP^2 = \frac{SS_1 + SS_2}{n_1 + n_2}$$

$$SS_1 = (14 - 10.875)^2 + (5 - 10.875)^2 + (4 - 10.875)^2 + (24 - 10.875)^2 + (8 - 10.875)^2 + (6 - 10.875)^2 + 2(13 - 10.875)^2$$

$$= 9.76 + 34.52 + 47.27 + 172.27 + 8.27 + 23.77 + 4.52$$

$$= 304.9$$

$$SS_2 = (19 - 16.17)^2 + (10 - 16.17)^2 + (32 - 16.17)^2 + (8 - 16.17)^2 + (15 - 16.17)^2 + (13 - 16.17)^2$$

$$= 8.01 + 38.07 + 250.59 + 66.75 + 1.37 + 10.05$$

$$= 374.84$$

$n_1 = 8 - 1 = 7$ $n_2 = 6 - 1 = 5$

$$SP^2 = \frac{304.9 + 374.84}{7 + 5} = 56.645$$

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{56.645}{8} + \frac{56.645}{6}} = 4.06$$

$$t = \frac{|10.875 - 16.17|}{4.06} = 1.3$$

$t_{0.05(2), 12} = 2.176 \rightarrow t \leq t_{0.05, 12}$ Fail to reject H_0

Notes from Lab 3

Single Sample T-test

- Example, Do ND students get 8 hrs sleep?

$$H_0: \bar{x}_{ND} = \mu \text{ (8 hrs)}$$

$$H_A: \bar{x}_{ND} \neq \mu$$

Independent T-test

- Comparing means b/w groups

$$H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 \neq \mu_2$$

Paired T-test

$$H_0: d = 0 \text{ (no diff)}$$

$$H_A: d \neq 0 \text{ (diff)}$$

P-values

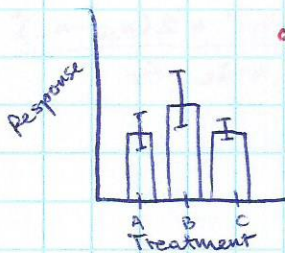
≤ 0.05 - Reject H_0

≥ 0.05 - Fail to reject

1/30/2014

Error

1. Can express as $\bar{x} \pm S$



Error bars can be expressed as SD

Shows population variation

- Always bigger than $S_{\bar{x}}$

2. Can express as $\bar{x} \pm S_{\bar{x}}$

- Tells you about the error in the mean.

- Sometimes expressed as $\bar{x} \pm 2S_{\bar{x}}$

- Always smaller than SD

3. Can express as $\bar{x} \pm t_{\alpha, \nu} S_{\bar{x}}$

- confidence interval mean falls within a range

Always give n value (number of experiments, measurements, etc.)

Hypothesis Test

- Statistical inference - drawing conclusion from a hypothesis test

- Start by constructing a null hypothesis H_0

↳ pneumonia example from last week $n = 1, 3, 5, 4, 7$ $\mu_0 = 7$ days to recover

For example:

$$\bar{x} = 4$$

w/o a treatment

$$H_0: \bar{x} = \mu \text{ or } \bar{x} = 7 \text{ days}$$

$$H_A: \bar{x} \neq \mu \text{ or } \bar{x} \neq 7 \text{ days}$$

1/30/2014

- Next we attempt to disprove null hypothesis so we can infer a result

Actual State of Nature

	H_0 True	H_0 False
Accept H_0	Correct $= 1 - \alpha$	Type II Error $= \beta$
Reject H_0	Type I Error $= \alpha$	Correct $1 - \beta = \text{"power"}$

$\alpha = P(\text{reject } H_0 | H_0 \text{ true})$
 $\beta = P(\text{accept } H_0 | H_0 \text{ false})$
5% is usually acceptable for likelihood of getting Type I Error

Meaning of Significance

Statistical significance: an apparent difference is likely not due to chance alone

Common application of hypothesis testing: t-test

↳ Compare two means

$H_0: \mu_1 = \mu_2$ Independent t-test (means are not overlapping)
 $H_A: \mu_1 \neq \mu_2$

Test statistic for above using t-dist

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{S_{\bar{x}_1 - \bar{x}_2}}$$

Standard Error of difference b/w means

$$= \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

$s_p^2 = \text{pooled variance}$

$$\hookrightarrow s_p^2 = \frac{SS_1 + SS_2}{n_1 + n_2 - 2} = \frac{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

- Example on green handout for today

The Significance Level α

- Common α -levels are 0.05 or 0.01

- When H_0 is rejected a certain α , we say:

the samples are significantly different at a probability $p \leq \alpha$

(where $\alpha = 0.05, 0.01, \text{etc.}$)

or more simply at $p = ?$

The probability value p

p or p -value which is the probability of observing more extreme value of the test statistic by chance alone

- we set α a priori then state specific p -value

- In science, "significant" should be reserved for significant statistical differences

Independent Groups t-test
Ex: t-test comparing samples of unequal size 1/30/2014

X_i = weight gain (g) of σ^7 rats raised on 2 different diets

	<u>1 = Low Protein Diet</u>	<u>2 = High Protein Diet</u>
	70	134
	118	107
	101	146
	85	83
	107	104
	132	119
	94	124
		161
		123
ΣX_i	707	1440
n	7	12
\bar{X}	101	120
to calculate pooled variance	$\left\{ \begin{array}{l} \Sigma_i (X_i - \bar{X})^2 \\ \nu \end{array} \right. \begin{array}{l} 2,552 \\ 6 \end{array}$	$\begin{array}{l} 5,032 \\ 11 \end{array}$

pooled $S^2 = \frac{2552 + 5032}{6 + 11} = 446$ with $\nu = 17$ $\leftarrow n_1 + n_2 - 2$

Null hypothesis
 $H_0: (\mu_1 - \mu_2) = (\bar{X}_1 - \bar{X}_2) = \emptyset$ or simply $H_0: \mu_1 = \mu_2$

HA: $\mu_1 \neq \mu_2$

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} = \frac{|101 - 120|}{\sqrt{\frac{446}{7} + \frac{446}{12}}} = \frac{19}{10} = 1.9$$

"test stat" \rightarrow two means differ by 1.9g
 used to compare to $t_{\alpha, \nu}$
 "critical value"

From t-table, $t_{0.05, 17} = 2.11$ (by interpolation, $p \approx 0.08$)
 α ν

$t \leq t_{0.05, 17}$ can't so reject H_0

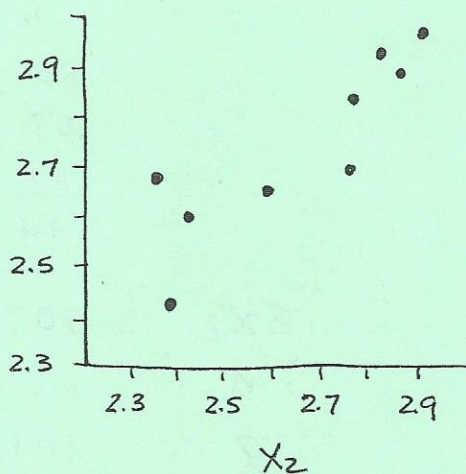
\therefore fail to reject H_0 ; diets are not significantly different

Paired t-test

Ex: TEST OF DIFFERENCE (Paired Comparison)

YEAR	X_1 POP. A	X_2 POP. B	ΔX $X_1 - X_2$ d	
1916	2.68	2.36	+0.32	Block 1
1917	2.60	2.41	+0.19	Block 2
1918	2.43	2.39	+0.04	↓ etc.
1919	2.90	2.85	+0.05	
1920	2.94	2.82	+0.12	
1921	2.70	2.73	-0.03	X_1
1922	2.66	2.58	+0.08	
1923	2.98	2.89	+0.09	
1924	2.85	2.78	+0.07	
mean	2.749	2.646	+0.103 = \bar{d}	

Association



t-test of $H_0: \mu_1 = \mu_2$ (Independent t-test)

$\nu = 2(9-1)$
↓
2 in each pop
critical $t_{0.05, 16} = 2.120$

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} = \frac{0.103}{0.093} = 1.11; \quad \nu = 16, p > 0.20$$

Fail to Reject H_0

parametric → paired differences = 0
t-test of $H_0: D = 0$ (Paired t-test)

↳ random years, assume parks exposed to same conditions and are paired

↳ use $x_1 - x_2$ measurements

$$t = \frac{\bar{d} - 0}{S_{\bar{d}}} = \frac{0.103}{0.0336} = 3.08; \quad \nu = 8, p < 0.02$$

$$S_{\bar{d}} = \frac{S}{\sqrt{n}} \text{ using } \Delta X \text{ data}$$

$n = 9$ observations
 $\nu = 8$

Can Reject H_0

1/30/2014

Biostatistics - BIOS 411
Definitions: Statistical Hypothesis Testing

Statistical hypothesis: A statement concerning the probability distribution of a random variable

Null hypothesis (H_0): A statement concerning the probability distribution that is to be tested, generally indicating that samples come from the same distribution

Alternative hypothesis (H_1): A statement concerning possible alternative probability distributions to the null hypothesis

Significance level (α -level): An error rate set *a priori* by the investigator, which is measured in terms of the probability of rejecting the null hypothesis when the H_0 is true

Probability value (p-value, or p-level): The probability of observing a more extreme value of the test statistic, than the value actually observed, by chance alone (i.e., given that the H_0 is true)

Test statistic: A computed value that provides a basis for assessing the null hypothesis in terms of a probability statement

Critical region: A set of values of the test statistic that lead to rejection of the null hypothesis

Critical value: The value that determines the borderline(s) of the critical region(s), given the α -level and the degrees of freedom

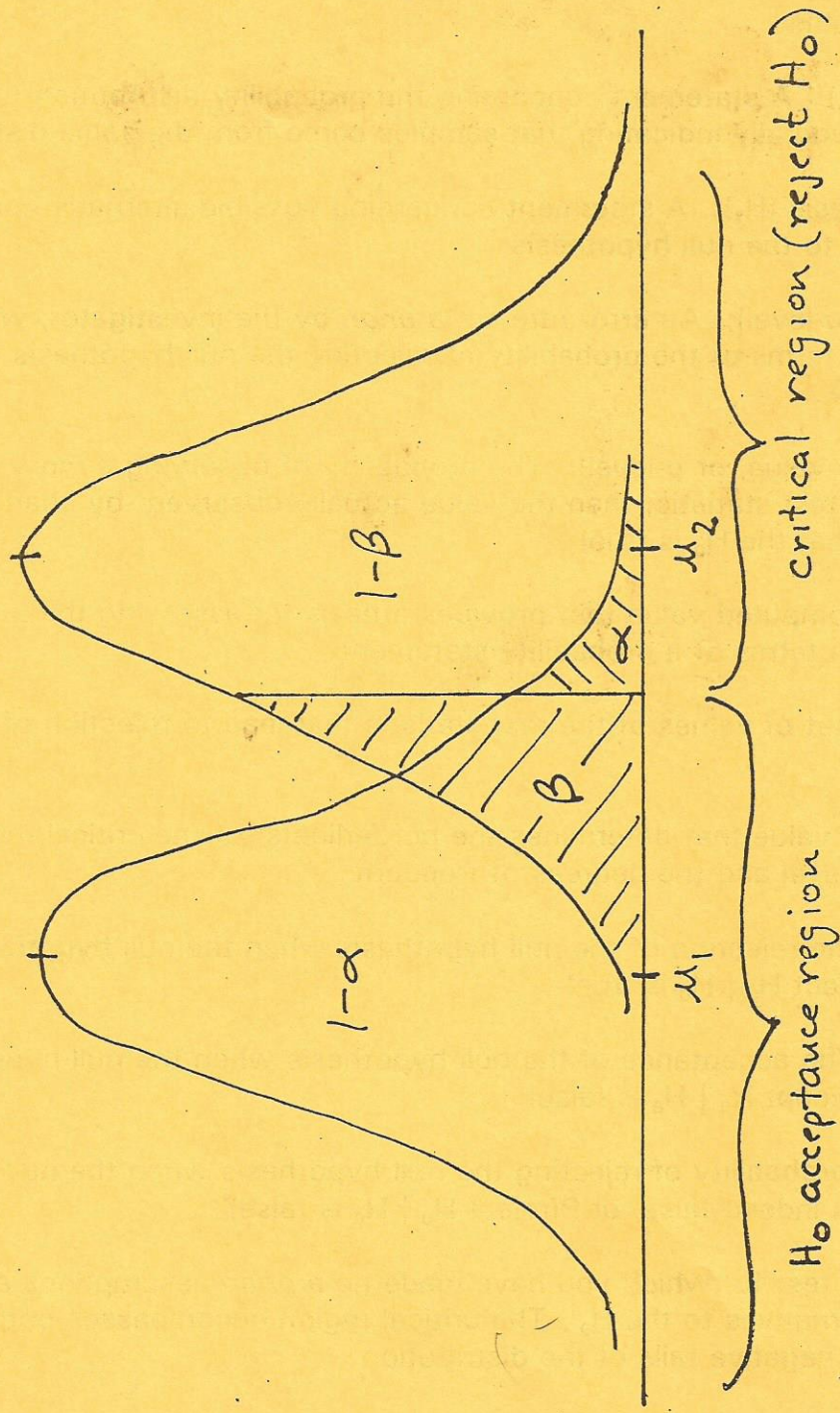
Type I error (α): The rejection of the null hypothesis when the null hypothesis is true, or $P(\text{reject } H_0 \mid H_0 \text{ is true})$

Type II error (β): The acceptance of the null hypothesis when the null hypothesis is false, or $P(\text{accept } H_0 \mid H_0 \text{ is false})$

Power ($1-\beta$): The probability of rejecting the null hypothesis when the null hypothesis is indeed false, or $P(\text{reject } H_0 \mid H_0 \text{ is false})$

Two-tailed test: A test for which you have made no *a priori* assumptions about possible alternatives to the H_0 . The critical region encompasses both positive and negative tails of the distribution

One-tailed test: A test for which you have stated a specific *a priori* alternative to the H_0 . The alternative hypothesis may be either larger or smaller than the H_0 , which corresponds to either positive or negative extreme values of the test statistic



1/30/2014

Mean of Significance Cent'd

- Biological significance means that the difference is large enough to be biologically relevant or important

How do we minimize β ?

- β = probability of Type II Error
- choose μ_1 far from μ_2 in experiment
- Accept larger Type I error (α)
- $\downarrow \beta$ but now have higher TI error...
- increase n to reduce tail error
 - ↳ more df

How many tails for a test?

- Two-tailed test means you have made no a priori assumptions about possible alternatives to H_0

$$\text{Two-tailed} \rightarrow H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 \neq \mu_2$$

↳ We don't know possible outcomes so we consider both tails

- One-tailed test means you have reason to believe that a sample pop can only be larger or smaller than that of the H_0 or you are only interested in one outcome

$$\text{One-tailed} \rightarrow H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 > \mu_2 \text{ (or } \mu_1 < \mu_2)$$

Example - drug testing

Tests of Association

- We may want to know about the association (or difference) between 2 variables
- One example is a paired t -test
- Blocking can be useful
 - Block = set of observations made under similar conditions
 - statistically remove the effects of block-to-block variation

2/4/2014

- Go online find pneumonia problem (sent 1131?)

Tests of Association

- We may want to know the association (or difference) b/w 2 variables
 - ↳ Example = paired t-test
- Blocking can be useful, can remove variation b/w paired observations
 - ↳ Block is a set of observations made under same conditions

- Classic example of paired test is before/after

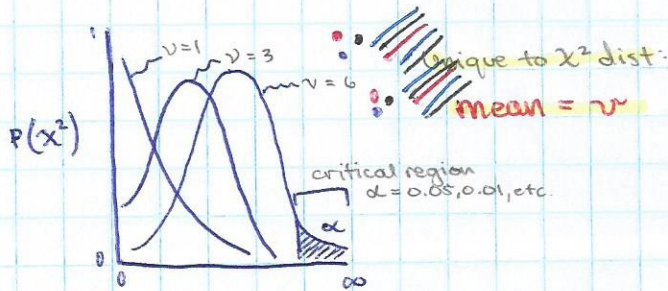
The χ^2 distribution

- Generally used for categorical data
- Compare observed set of values to expected set (under the H_0)
- We can test if two variables X_1, X_2 are independent (H_0) or are somehow dependent on each other
- Goodness of fit test allows us to compare observed and expected distribution

$$\chi^2 = \sum_{i=1}^k \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

$$= \sum_{i=1}^k \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} \quad (\text{Zar equation})$$

k = categories f_i = observed frequency \hat{f}_i = expected frequency



• Shape highly dependent on degrees of freedom

Contingency Tables

- Expectations contingent on observations
- Used to summarize categorical data for χ^2 analysis of ≥ 2 variables
- Example

	A	B
Gender X_1	p	1-p
Hats? X_2	q	1-q

H_0 : independence of gender and hats

H_A : not independent

• Expectations arrive from multiplication of independent $P(X)$

	A	B
X_1		
A	pq	q(1-p)
B	p(1-q)	(1-p)(1-q)

Expectations under H_0

2/4/2014

Chi-Square Contingency Table

Suppose that we have two variables, sex (male or female) and handedness (right- or left-handed). Further suppose that 100 individuals are randomly sampled from a very large population as part of a study of sex differences in handedness. A contingency table below displays the numbers of individuals who are male and right-handed, male and left-handed, female and right-handed, and female and left-handed.

Test the null hypothesis that handedness is independent of gender.

	Right-handed	Left-handed	Total
Males	43	9	52
Females	44	4	48
Totals	87	13	100

$$exp_{M,R} = \frac{52}{100} \times \frac{87}{100} \times 100 = 45.24 \quad exp_{M,L} = \frac{52}{100} \times \frac{13}{100} \times 100 = 6.76$$

$$exp_{F,R} = \frac{48}{100} \times \frac{87}{100} \times 100 = 41.76 \quad exp_{F,L} = \frac{48}{100} \times \frac{13}{100} \times 100 = 6.24$$

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(43-45.24)^2}{45.24} + \frac{(9-6.76)^2}{6.76} + \frac{(44-41.76)^2}{41.76} + \frac{(4-6.24)^2}{6.24}$$

$$= .111 + 0.74 + .12 + .80$$

$$\chi^2 = 1.77$$

$$\chi^2_{0.05,1} = 3.841$$

$\chi^2 \leq \chi^2_{0.05,1} \rightarrow$ Can't reject H_0 , handedness independent of gender

$$0.1 < p < 0.5$$

~~So results not significant and may reject H_0 ?~~

No snips 'n snails, boys linked to PCBs

By **WAYNE FALDA**
Tribune Staff Writer

Sons of Lake Michigan fishermen may owe their gender to high levels of PCBs in the fish their fathers eat, according to a recently published study by Michigan State University researchers.

Epidemiologist Wilfried Karmaus led a study that examined the levels of PCBs in the blood of fathers of 101 families and discovered that 57 percent of the 208 children were males.

Birth studies have shown that slightly more females are born than males on average across the U.S. populations.

"If there was a female effect, it would be easier to explain. But we had a male effect."

Wilfried Karmaus
Epidemiologist

Scientists have conducted many studies intended to see whether environmental contaminants would affect human reproduction.

The conclusions of these various studies have produced mixed results.

See PCBS/Page A7

/JIM RIDER

idrew,
after

Question: Was there an equal number of male and female children among the 208 offspring?

Offspring = 208
Males = $.57 \times 208 = 118.56$ $H_0: M:F = 1:1$
Females = $.43 \times 208 = 89.44$

K	O	E	$(O-E)^2/E$
M	118.56	104	2.0384
F	89.44	104	2.0384

$\chi^2 = 4.0768$ $\chi_{0.05,1} = 3.841$

$\chi^2 \geq \chi_{0.05,1}^2$

Reject $H_0 \rightarrow$ Not equal number of children

χ^2 Contingency Table (Test for Independence)

2/4/2014

Observations:

		Wild Celery (WC)		TOTAL	
		Absent	Present		
Eurasian Water Milfoil (EWM)	Absent	1 (19.3)	41 (22.7)	42	G test $1 \ln\left(\frac{1}{19.3}\right) + 41 \ln\left(\frac{41}{22.7}\right) + 45 \ln\left(\frac{45}{26.7}\right) + 13 \ln\left(\frac{13}{31.3}\right) = 33.34$ $\times 2 =$ $G = 66.69$
	Present	45 (26.7)	13 (31.3)	58	
	TOTAL	46	54	100	

Multiplication of independent probabilities:

P(absent)		P(present)		Expectations under H_0 : independence	
				WC	
		Absent	Present		
WC	.46	.54			
EWM	.42	.58	Absent	$.46 \times .42 \times 100 = 19.3$	$.42 \times .54 \times 100 = 22.7$
			Present	$.46 \times .58 \times 100 = 26.7$	$.54 \times .58 \times 100 = 31.3$

Efficient approach using marginals:

1	41	42	expectation for this cell =
			$\frac{42}{100} \times \frac{54}{100} \times 100 = 22.7$
45	13	58	G test
46	54	100	$p < 0.001$ Greater

Compute χ^2 and compare to critical χ^2 :

$$\chi^2 = \sum \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = \frac{(1-19.3)^2}{19.3} + \frac{(45-26.7)^2}{26.7} + \frac{(41-22.7)^2}{22.7} + \frac{(13-31.3)^2}{31.3} = 55.5$$

$$\text{crit. } \chi^2_{\alpha, \nu} = \chi^2_{0.05, 1} = 3.841$$

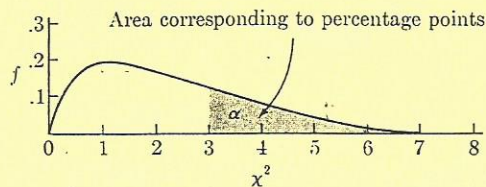
Conclusion: Reject H_0 that WC and EWM occur independently. They are more likely to occur separately than together.

Biological reason: EWM (an exotic) overgrows WC in deep water. In shallow water, WC survives ice but EWM does not.

TABLE IV
Critical values of the chi-square distribution

ν	α										ν
	.995	.975	.9	.5	.1	.05	.025	.01	.005	.001	
1	0.000	0.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	10.828	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	13.816	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	16.266	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	18.467	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	20.515	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	22.458	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	24.322	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	26.124	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	27.877	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	29.588	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	31.264	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	32.910	12
13	3.565	5.009	7.042	12.340	19.812	22.362	24.736	27.688	29.819	34.528	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	36.123	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	37.697	15
16	5.142	6.908	9.312	15.338	23.542	26.296	28.845	32.000	34.267	39.252	16
17	5.697	7.564	10.085	16.338	24.769	27.587	30.191	33.409	35.718	40.790	17
18	6.265	8.231	10.865	17.338	25.989	28.869	31.526	34.805	37.156	42.312	18
19	6.844	8.907	11.651	18.338	27.204	30.144	32.852	36.191	38.582	43.820	19
20	7.434	9.591	12.443	19.337	28.412	31.410	34.170	37.566	39.997	45.315	20
21	8.034	10.283	13.240	20.337	29.615	32.670	35.479	38.932	41.401	46.797	21
22	8.643	10.982	14.042	21.337	30.813	33.924	36.781	40.289	42.796	48.268	22
23	9.260	11.688	14.848	22.337	32.007	35.172	38.076	41.638	44.181	49.728	23
24	9.886	12.401	15.659	23.337	33.196	36.415	39.364	42.980	45.558	51.179	24
25	10.520	13.120	16.473	24.337	34.382	37.652	40.646	44.314	46.928	52.620	25
26	11.160	13.844	17.292	25.336	35.563	38.885	41.923	45.642	48.290	54.052	26
27	11.808	14.573	18.114	26.336	36.741	40.113	43.194	46.963	49.645	55.476	27
28	12.461	15.308	18.939	27.336	37.916	41.337	44.461	48.278	50.993	56.892	28
29	13.121	16.047	19.768	28.336	39.088	42.557	45.722	49.588	52.336	58.301	29
30	13.787	16.791	20.599	29.336	40.256	43.773	46.979	50.892	53.672	59.703	30
31	14.458	17.539	21.434	30.336	41.422	44.985	48.232	52.191	55.003	61.098	31
32	15.134	18.291	22.271	31.336	42.585	46.194	49.480	53.486	56.329	62.487	32
33	15.815	19.047	23.110	32.336	43.745	47.400	50.725	54.776	57.649	63.870	33
34	16.501	19.806	23.952	33.336	44.903	48.602	51.966	56.061	58.964	65.247	34
35	17.192	20.569	24.797	34.336	46.059	49.802	53.203	57.342	60.275	66.619	35
36	17.887	21.336	25.643	35.336	47.212	50.998	54.437	58.619	61.582	67.985	36
37	18.586	22.106	26.492	36.335	48.363	52.192	55.668	59.892	62.884	69.346	37
38	19.289	22.878	27.343	37.335	49.513	53.384	56.896	61.162	64.182	70.703	38
39	19.996	23.654	28.196	38.335	50.660	54.572	58.120	62.428	65.476	72.055	39
40	20.707	24.433	29.051	39.335	51.805	55.758	59.342	63.691	66.766	73.402	40
41	21.421	25.215	29.907	40.335	52.949	56.942	60.561	64.950	68.053	74.745	41
42	22.138	25.999	30.765	41.335	54.090	58.124	61.777	66.206	69.336	76.084	42
43	22.859	26.785	31.625	42.335	55.230	59.304	62.990	67.459	70.616	77.419	43
44	23.584	27.575	32.487	43.335	56.369	60.481	64.202	68.710	71.893	78.750	44
45	24.311	28.366	33.350	44.335	57.505	61.656	65.410	69.957	73.166	80.077	45
46	25.042	29.160	34.215	45.335	58.641	62.830	66.617	71.201	74.437	81.400	46
47	25.775	29.956	35.081	46.335	59.774	64.001	67.821	72.443	75.704	82.720	47
48	26.511	30.755	35.949	47.335	60.907	65.171	69.023	73.683	76.969	84.037	48
49	27.249	31.555	36.818	48.335	62.038	66.339	70.222	74.919	78.231	85.351	49
50	27.991	32.357	37.689	49.335	63.167	67.505	71.420	76.154	79.490	86.661	50

Note: For values of $\nu > 100$, compute approximate critical values of χ^2 by formula as follows: $\chi^2_{\alpha(\nu)} = \frac{1}{2}(t_{2\alpha(\nu)} + \sqrt{2\nu - 1})^2$, where $t_{2\alpha(\nu)}$ can be looked up in Table III. Thus $\chi^2_{0.05(120)}$ is computed as $\frac{1}{2}(t_{0.10(120)} + \sqrt{240 - 1})^2 = \frac{1}{2}(1.645 + \sqrt{239})^2 = \frac{1}{2}(17.10462)^2 = 146.284$. For $\alpha > 0.5$ employ $t_{1-2\alpha(\nu)}$ in the above formula. When $\alpha = 0.5$, $t_{2\alpha} = 0$. Values of chi-square from 1 to 30 degrees of freedom have been taken from a more extensive table by C. M. Thompson (*Biometrika* 32:188-189, 1941) with permission of the publisher.



2/4/2014

- For large RxC tables, use computers!
- Slightly more powerful is the G-test (will talk about later)

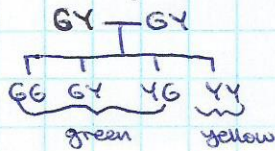
χ^2 goodness of fit to an Expected Distribution

Two ways to use χ^2

1. Internal expected distribution
 - calculated from data, ex. contingency tables
2. External expected distribution
 - Familiar biological examples \rightarrow Mendelian genetics
 - \hookrightarrow known beforehand; used to make H_0

Examples: Peas

\hookrightarrow green dominant to yellow



H_0 : observed ratio = expected ratio of 3:1

122 off spring of unknown cross

\hookrightarrow 98 green and 24 yellow

Outcome	Obs	exp
G	98	91.5
Y	24	30.5

} calculated external ratio based on 3:1 law

$$\frac{(\text{obs} - \text{exp})^2}{\text{exp}}$$

$$0.462$$

$$1.385$$

$$\hat{\chi}^2 = 1.847$$

$$df = k - 1 = 2 - 1 = 1$$

two pea colors

$$\text{crit } \chi^2_{0.05, 1} = 3.841$$

$$\hat{\chi}^2 \leq \chi^2_{0.05, 1} \rightarrow \text{Do not reject } H_0$$

2/6/2014

χ^2 Testing Proportions

Ex.

100 seeds \rightarrow expose to 10°C 100 seeds \rightarrow 15°C

H_0 : Germination independent of T

		Germinated		
		N	Y	
Temp	10°	a	b	100
	15°	c	d	100
		a+c	b+d	200

As t-test:

H_0 : % germ at 10° = % germ at 15°

$$t = \frac{|\bar{X}_{10} - \bar{X}_{15}|}{\sqrt{\frac{SE^2}{n} + \frac{SE^2}{n}}}$$

\hookrightarrow Need at least 2 data sets

- χ^2 test has less power

2/6/2014

Low frequencies make χ^2 table less accurate

strongly influenced by values < 5

Alternative test is Fisher Exact

could increase sample size or divide batch so $f \geq 5$

G-test for Analyzing Frequency

improved method for Goodness of Fit

better for great departures between obs and exp values

- If $|f_i - \hat{f}_i| > \hat{f}_i$ probably don't want to do χ^2

- G test

$G = 2 \ln L$ still use χ^2 table for value of $G_{\alpha, r}$

↳ what is L?

$$\ln L = \sum_{i=1}^k f_{obs} \ln \left(\frac{f_{obs}}{f_{exp}} \right)$$

also called log-likelihood test

↳ In subdivides large deviations b/w obs and exp

Written also as: $G = 2 \sum_{i=1}^k f_{obs} \ln \left(\frac{f_{obs}}{f_{exp}} \right)$

Comparing χ^2 and G

About equivalent at large N

Differences occur at smaller N

For both, small frequencies should be combined if possible

α is also a bit higher for G

• More likely to reject true H_0

END MATERIAL EXAM I

Exam is 2/3 problems, 1/3 definitions, short answers

↳ can use calculator

↳ 3x5 index card w/ formulas (one-side), no words only formulas

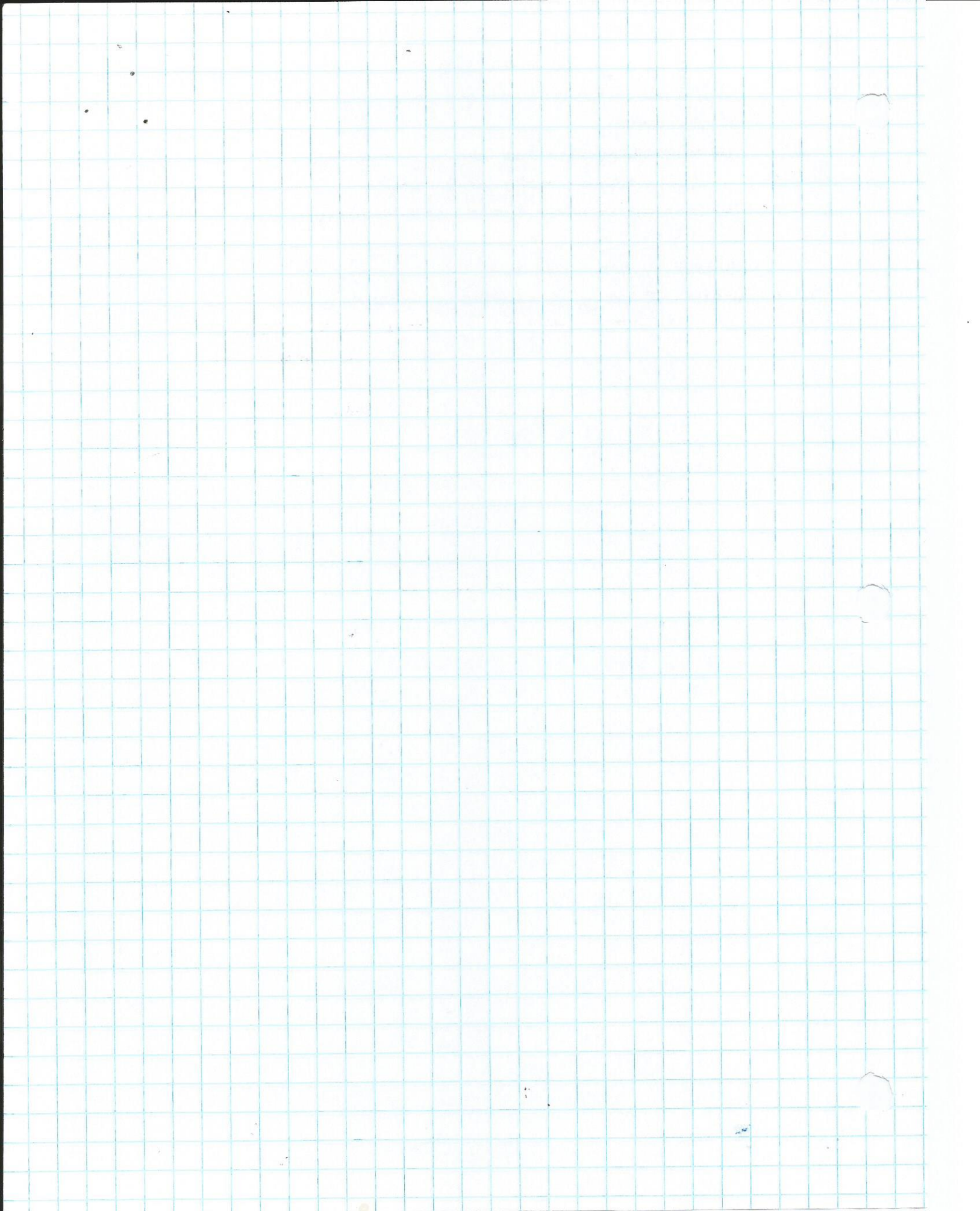
Recap section I

1. Basic info - variables, populations, samples
2. Descriptive stats - location, dispersion
3. Probability distributions - discrete, continuous
discrete - binomials
4. Continuous distributions
 - (1) Normal distribution (Z)
 - (2) T-distribution (for smaller samples)
 - (a) attributes
 - (b) confidence limits (95% CI)
 - (c) \bar{x} vs. μ (pneumonia patients)

2/6/2014

- (d) \bar{x}_1 vs. \bar{x}_2 (rat protein diet)
- (e) \bar{d} vs. 0 (wolf births per year)
- (3) χ^2 distribution
 - (a) Internal expected dist (obs to generate exp)
 - (i) contingency tables (pond weeds)
 - (b) External expected dist (obs to external exp)
 - (i) Genetics ratios
- (4) G-test as alternative to χ^2

Review 6-8 PM 143 Ecology lab in Jordan on Wednesday



Exam 2

2/11/2014

Analysis of Variance (ANOVA)

• A. Fisher - important figure in statistics

• One application of test is testing whether 2 or more sample means come from the same population with mean μ

Ex. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

Can we adjust α -level for multiple t-tests?

- see blue pg

$$\alpha_{\text{corr}} = \frac{0.05}{\# \text{ tests}} = \frac{0.05}{6} = 0.0083$$



Bonferroni correction

↳ Run each test at lower α to compensate

- Raises type II errors tho...



Power declines if we compensate so we could increase n to increase the power

• but as # of treatments increase the required n increases proportional to the square of treatments

- So compensation not best \rightarrow ANOVA comes in to save the day!

Deductions in ANOVA

Deduction 1 If H_0 is true then variances among replicates within groups estimates the population variance (σ)

S_W^2 within groups variance (pooled σ among replicates w/n groups)

$$S_W^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$$

k = all groups n = replicates/group

Double summation (sum all replicates, then all groups of replicates)

$k(n-1) = \nu_2$ degrees of freedom for ANOVA

Example on blue sheet $\nu_2 = (4)(6-1) = 20$

• Recall T-test

$$S_p^2 = \frac{SS_1 + SS_2}{\nu_1 + \nu_2} \rightarrow \text{Do this for each group essentially}$$

↳ simply an extension of t-test

Deduction 2 If H_0 is true then variance among treatments also estimates the population variance

• S_A^2 = variance among treatments; "among-groups variance"

$$S_A^2 = \frac{n}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2$$

n = for balanced treatments, put on even playing field

ν_1 = among groups df \rightarrow Example = $4-1 = 3$ df



each SS is based on n replicates

2/11/2014

Deduction 3 if H_0 true, then within groups variance and among group

variance estimate the same thing (pop. variance)

- $S_W^2 = S_A^2$ if H_0 true
- $F = \frac{S_A^2}{S_W^2} = 1$ if H_0 true

The F-statistic has two dfs (ν_1, ν_2)

For donut prob on blue pg:

$$F = \frac{S_A^2}{S_W^2} = \frac{545.5}{100.9} = 5.4 \text{ with } \nu_1 = 3 \text{ } \nu_2 = 20$$

df=3 df=20

two ν are necessary for reading F-distribution table (blue pg)

For $F_{\alpha, \nu_1, \nu_2} = F_{0.05, 3, 20} = 3.10$

$F_{0.05, 3, 20} \rightarrow$ Reject H_0

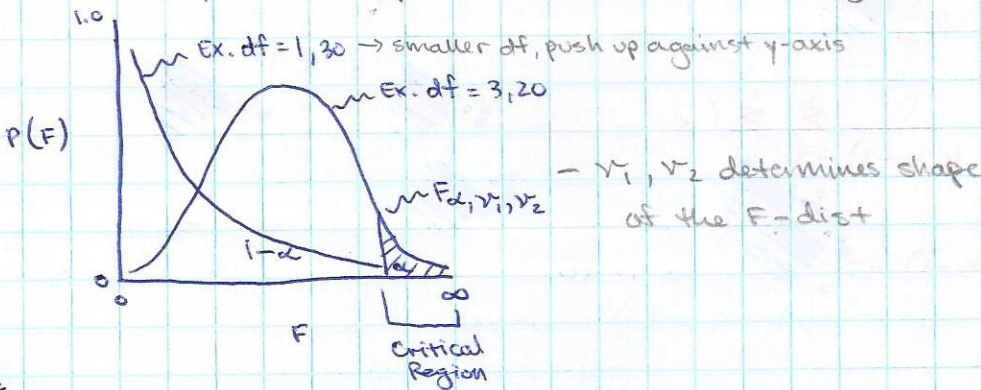
Variance among groups is greater than expected by chance alone, so unlikely ($p < 0.01$) H_0 is true

this was a one way ANOVA or single factor ANOVA

↳ one variable for donuts = oil type

Summary points for F-test

- F-stat is calculated from 2 variances, the among and within variance
- The dist. of repeated samplings and calculation of F yields F-distribution



For $k=2$ groups, the F-test is mathematically equivalent to t-test

- For F-test ANOVA

$\nu_1 = 1, \nu_2 = \infty$ (infinite replicates)

↳ $F(1, \infty) = t_{\infty}^2$ $F(1, \infty) = 3.84$ $t_{0.05, \infty} = 1.96 \rightarrow t^2 = 3.84$ OMG

* Linear Models

$X_{obs.} =$ biological mechanism + summation of error

Parametric Linear Model

$$x_{ij} = \mu + \tau_i + \epsilon_{ij}$$

τ_i = treatment effect

μ = parametric grand mean

ϵ_{ij} = all natural variation or error

Estimated Model

$$\hat{x}_{ij} = \bar{\bar{x}} + (\bar{x}_i - \bar{\bar{x}}) + (x_{ij} - \bar{x}_i)$$

$\bar{\bar{x}}$ = estimated grand mean

$(\bar{x}_i - \bar{\bar{x}})$ = estimated treatment effect

$(x_{ij} - \bar{x}_i)$ = estimate of error

ONE-WAY ANALYSIS OF VARIANCE (ANOVA), or F-TEST

2/11/2014

Example: Cook doughnuts in 4 kinds of fat (treatments);
 X_i = grams of fat absorbed by 24 doughnuts (response)

		Treatments ($k=4$)				
		Land	Peanut Oil	corn Oil	Safflower Oil	
		78	75	64	55	
		91	93	72	66	replicate - individual unit to which a specific treatment has been applied - replication - multiple units exposed to specific treatment
balanced design = equal number of replicants per group (b) ↳ Good to have equal no.		97	78	68	49	
		82	71	77	64	
		85	63	56	70	
		77	76	95	68	
	\bar{x}	85	76	72	62	

$$H_0: \mu_{\text{land}} = \mu_{\text{po.}} = \mu_{\text{c.o.}} = \mu_{\text{s.o.}}$$

What could we do to compare these? - Ex. pairwise t-tests for k treatments

$$\# \text{tests} = k(k-1)/2 = (4)(4-1)/2 = 6 \text{ tests to compare all samples w/ t-test}$$

say $\alpha = 0.05$ for any test

$$\text{Expected type I error} = \alpha \times \# \text{tests} = 0.05 \times 6 = .30 \text{ "higher chance of Type I Error}$$

↳ procedure wise error rate - compounding type I error

H_A : all means are not equal

- might be varying degrees of difference; this H_A accounts for all possible combos

		Treatments (i)			
		1	2	3	4
Replicates (j)	1	$(x_{11} - \bar{x}_1)^2$	$(x_{21} - \bar{x}_2)^2$		
	2	$(x_{12} - \bar{x}_1)^2$	$(x_{22} - \bar{x}_2)^2$		
	3	$(x_{13} - \bar{x}_1)^2$			
	4	$(x_{14} - \bar{x}_1)^2$			
	5	$(x_{15} - \bar{x}_1)^2$			
	6	$(x_{16} - \bar{x}_1)^2$			
	\bar{x}_i	$(\bar{x}_1 - \bar{\bar{x}})^2$	$(\bar{x}_2 - \bar{\bar{x}})^2$	$(\bar{x}_3 - \bar{\bar{x}})^2$	$(\bar{x}_4 - \bar{\bar{x}})^2$

REMEMBER!
 "ith group
 jth replicate"

ANOVA

TABLE V
Critical values of the F distribution

		ν_1 (degrees of freedom of numerator mean squares)												
		1	2	3	4	5	6	7	8	9	10	11	12	α
ν_2 (degrees of freedom of denominator mean squares)	1 .05	161	199	216	225	230	234	237	239	241	241	243	244	.05
	.025	648	800	864	900	922	937	948	957	963	969	973	977	.025
	.01	4050	5000	5400	5620	5760	5860	5930	5980	6020	6060	6080	6110	.01
	2 .05	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	.05
	.025	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	.025
	.01	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	.01
	3 .05	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	.05
	.025	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.3	.025
	.01	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	27.1	.01
	4 .05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	.05
	.025	12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.79	8.75	.025
	.01	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.4	.01
5 .05	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.71	4.68	.05	
.025	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.57	6.52	.025	
.01	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.99	9.89	.01	
6 .05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	.05	
.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.41	5.37	.025	
.01	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	.01	
7 .05	5.59	4.74	4.35	4.12	3.97	3.87	3.77	3.73	3.68	3.64	3.60	3.57	.05	
.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.89	4.82	4.76	4.71	4.67	.025	
.01	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47	.01	
8 .05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	.05	
.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.25	4.20	.025	
.01	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67	.01	
9 .05	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	.05	
.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.91	3.87	.025	
.01	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11	.01	
10 .05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	.05	
.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.67	3.62	.025	
.01	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71	.01	

		1	2	3	4	5	6	7	8	9	10	11	12	α
ν_2 (degrees of freedom of denominator mean squares)	11 .05	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	.05
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.48	3.43	.025
	.01	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40	.01
	12 .05	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	.05
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.32	3.28	.025
	.01	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16	.01
	15 .05	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	.05
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	3.01	2.96	.025
	.01	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	.01
	20 .05	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	.05
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.72	2.68	.025
	.01	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23	.01
24 .05	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	.05	
.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.59	2.54	.025	
.01	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03	.01	
30 .05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	.05	
.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.46	2.41	.025	
.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.90	2.84	.01	
40 .05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.04	.05	
.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.33	2.29	.025	
.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66	.01	
60 .05	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	.05	
.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.22	2.17	.025	
.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	.01	
120 .05	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83	.05	
.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.10	2.05	.025	
.01	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34	.01	
∞ .05	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75	.05	
.025	5.02	3.69	3.11	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.99	1.94	.025	
.01	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18	.01	

2/18/2014

Linear Additive Models cont'd

Rearranged Estimated Model: $(x_{ij} - \bar{x}_i) = (\bar{x}_i - \bar{x}) + (x_{ij} - \bar{x}_i)$

↳ This partitions the variance between parameters

ANOVA Table - where $k = \#$ treatments, $n = \#$ of replicates per treatment (groups)

Source of Variation	df	SS	Mean Square (MS)	F	P
Among treatments (Groups)	$k-1$ (v_1)	$n \sum_{i=1}^k (\bar{x}_i - \bar{x})^2$ $\sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$	SS_A / df_A	MS_A / MS_W	?
Within treatments (Error, residual)	$k(n-1)$ (v_2) $\sum_{i=1}^k (n_i - 1)$	$\sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$ ↳ The n will vary per summation	SS_W / df_W	this to F-table	↑
Total Var.	$(kn) - 1$ $(\sum_{i=1}^k n_i) - 1$	$\sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$ ↳ n varies per summation			These totals should equal sum of parts

Design?
Balanced
Unbalanced
 $n_i = \#$ reps in specific groups (weighting factor)

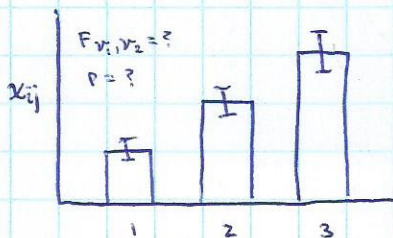
Back to donut problem

Source	df	SS	MS	F	P
Among	$v_1 = 3$	1636.5	545.5	5.4	< .01
Within	$v_2 = 20$	2018	100.9		
Total	23	3654.5			

You will never get anymore df than total # of replicates - 1

Rejecting H_0 in ANOVA

- ANOVA tests H_0 : all means are equal
- If we reject H_0 , means there is at least one difference somewhere
- ANOVA does not tell you which specific means are different



- Can just say the means are different and leave it at that or...

Use Multiple Comparison Test (MCT)

Post-hoc contrast test to see

where specific differences lie

Only do this w/ significant rejection of H_0

Multiple Comparison Test

"Multiple Contrasts"

Used only after getting significant ANOVA - rejection H_0

- Just like multiple t-tests, all MCTs must deal w/ the problem of controlling the effective α

2/18/2014

I. Types of MCTs

Two general categories

Planned comparisons: you identify specific contrasts you want to do

priori to ANOVA

Unplanned Comparisons: you identify the contrasts a posteriori to ANOVA

(normally you do all of them)

I. Planned Comparisons

Sometimes we know a subset of contrasts beforehand that we want to make

(ex. control vs. several treatments)

Least significant difference (Fisher's LSD) can be used

↳ Don't problem again...

		Lard (85)	Peanut (76)	Corn (72)
diff. of means	Soft (62)	23	14	10
	Compare			

LSD will actually compute minimum mean needed to call to means different

$$LSD = t_{\alpha(2), \nu_2} \sqrt{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

from ANOVA, two treatments being compared

For Donuts $LSD = 2.086 \sqrt{100.9 \left(\frac{1}{6} + \frac{1}{6} \right)} = 12.1 \rightarrow$ minimum difference

easier for balanced design

would need to do several of these for an unbalanced design

↳ $85 - 62 = 23$, $76 - 62 = 14$ - different means

2/20/14

- With LSD, in essence, we've done a t-test

$$t_s = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t_{\alpha(2), \nu_2}$$

Advantage of LSD vs. t-test is that MS_w is pooled error from ANOVA

but LSD is not very conservative - tend to find more differences than actually exist (α -level is not well controlled because it gets compounded)

II. Unplanned Comparisons

These contrasts are suggested by the data

Hsu's Highly Significant Difference (HSD) test

More conservative control of α ; may find fewer differences than actually exist $\rightarrow \beta$ is higher than desired

2/18/14

Tukey's HSD Table

App62

Statistical Tables and Graphs

Appendix B

TABLE B.5 (cont.) Critical Values of the q Distribution

$\alpha = 0.05$

v	$k(\text{or } p) = 2$	3	4	5	6	7	8	9	10
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07
2	6.085	8.331	9.798	10.88	11.74	12.44	13.03	13.54	13.99
3	4.501	5.910	6.825	7.502	8.037	8.478	8.853	9.177	9.462
4	3.927	5.040	5.757	6.287	6.707	7.053	7.347	7.602	7.826
5	3.635	4.602	5.218	5.673	6.033	6.330	6.582	6.802	6.995
6	3.461	4.339	4.896	5.305	5.628	5.895	6.122	6.319	6.493
7	3.344	4.165	4.681	5.060	5.359	5.606	5.815	5.998	6.158
8	3.261	4.041	4.529	4.886	5.167	5.399	5.597	5.767	5.918
9	3.199	3.949	4.415	4.756	5.024	5.244	5.432	5.595	5.739
10	3.151	3.877	4.327	4.654	4.912	5.124	5.305	5.461	5.599
11	3.113	3.820	4.256	4.574	4.823	5.028	5.202	5.353	5.487
12	3.082	3.773	4.199	4.508	4.751	4.950	5.119	5.265	5.395
13	3.055	3.735	4.151	4.453	4.690	4.885	5.049	5.192	5.318
14	3.033	3.702	4.111	4.407	4.639	4.829	4.990	5.131	5.254
15	3.014	3.674	4.076	4.367	4.595	4.782	4.940	5.077	5.198
16	2.998	3.649	4.046	4.333	4.557	4.741	4.897	5.031	5.150
17	2.984	3.628	4.020	4.303	4.524	4.705	4.858	4.991	5.108
18	2.971	3.609	3.997	4.277	4.495	4.673	4.824	4.956	5.071
19	2.960	3.593	3.977	4.253	4.469	4.645	4.794	4.924	5.038
20	2.950	3.578	3.958	4.232	4.445	4.620	4.768	4.896	5.008
24	2.919	3.532	3.901	4.166	4.373	4.541	4.684	4.807	4.915
30	2.888	3.486	3.845	4.102	4.302	4.464	4.602	4.720	4.824
40	2.858	3.442	3.791	4.039	4.232	4.389	4.521	4.635	4.735
60	2.829	3.399	3.737	3.977	4.163	4.314	4.441	4.550	4.646
120	2.800	3.356	3.685	3.917	4.096	4.241	4.363	4.468	4.560
∞	2.772	3.314	3.633	3.858	4.030	4.170	4.286	4.387	4.474

v	$k(\text{or } p) = 11$	12	13	14	15	16	17	18	19
1	50.59	51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83
2	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57
3	9.717	9.946	10.15	10.35	10.53	10.69	10.84	10.98	11.11
4	8.027	8.208	8.373	8.525	8.664	8.794	8.914	9.028	9.134
5	7.168	7.324	7.466	7.596	7.717	7.828	7.932	8.030	8.122
6	6.649	6.789	6.917	7.034	7.143	7.244	7.338	7.426	7.508
7	6.302	6.431	6.550	6.658	6.759	6.852	6.939	7.020	7.097
8	6.054	6.175	6.287	6.389	6.483	6.571	6.653	6.729	6.802
9	5.867	5.983	6.089	6.186	6.276	6.359	6.437	6.510	6.579
10	5.722	5.833	5.935	6.028	6.114	6.194	6.269	6.339	6.405
11	5.605	5.713	5.811	5.901	5.984	6.062	6.134	6.202	6.265
12	5.511	5.615	5.710	5.798	5.878	5.953	6.023	6.089	6.151
13	5.431	5.533	5.625	5.711	5.789	5.862	5.931	5.995	6.055
14	5.364	5.463	5.554	5.637	5.714	5.786	5.852	5.915	5.974
15	5.306	5.404	5.493	5.574	5.649	5.720	5.785	5.846	5.904
16	5.256	5.352	5.439	5.520	5.593	5.662	5.727	5.786	5.843
17	5.212	5.307	5.392	5.471	5.544	5.612	5.675	5.734	5.790
18	5.174	5.267	5.352	5.429	5.501	5.568	5.630	5.688	5.743
19	5.140	5.231	5.315	5.391	5.462	5.528	5.589	5.647	5.701
20	5.108	5.199	5.282	5.357	5.427	5.493	5.553	5.610	5.663
24	5.012	5.099	5.179	5.251	5.319	5.381	5.439	5.494	5.545
30	4.917	5.001	5.077	5.147	5.211	5.271	5.327	5.379	5.429
40	4.824	4.904	4.977	5.044	5.106	5.163	5.216	5.266	5.313
60	4.732	4.808	4.878	4.942	5.001	5.056	5.107	5.154	5.199
120	4.641	4.714	4.781	4.842	4.898	4.950	4.998	5.044	5.086
∞	4.552	4.622	4.685	4.743	4.796	4.845	4.891	4.934	4.974

Analysis of Variance Example

A manager wishes to determine whether the mean times required to complete a certain task differ for the three levels of employee training. He randomly selects 10 employees within each of the three levels of training (Beginner, Intermediate, and Advanced) and tests them in task completion (x_i = minutes to complete task). Do the data provide sufficient evidence to indicate that the mean times required to complete a certain task differ among the three levels of training? The data are summarized below, including the sums of squares.

Level of Training	n	\bar{x}	diff means	s^2
Advanced	10	24.2	} 2.9 } 3.1	21.54
Intermediate	10	27.1		6.0
Beginner	10	30.2		17.76

Null $H_0: \mu_A = \mu_I = \mu_B$ ✓
 $H_A: \mu_A \neq \mu_I \neq \mu_B$ ✓

$SS_W = 521.5$

$SS_A = 180.1$

$$S_W^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \quad S_A^2 = \frac{n}{k-1}$$

1. Perform the remainder of the problem, including choosing your alpha level, constructing null and alternative hypotheses, setting up your ANOVA table, and testing your null hypotheses. Draw a conclusion with a p-value.

2. If you reject the null hypothesis, then choose and perform a multiple comparisons test and draw your conclusion(s). Justify your choice of MCT.

Source of Variation	df	SS	Mean Sq	F	P
Among Treatments	2	180.1	90.05	4.66	$\alpha < 0.05$
Within Treatments	27	521.5	19.31	$F_{0.05, 2, 27} = 3.35$	
Total	29	701.6		$F > 3.35$	Reject H_0

$LSD = t_{\alpha(2), \nu_2} \sqrt{MS_W (\frac{1}{n_1} + \frac{1}{n_2})}$ $t_{0.05(2), 27} = 2.052$

Tukeys = $q_{\alpha(2), \nu_2} \sqrt{\frac{MS_W}{2} (\frac{1}{n_1} + \frac{1}{n_2})}$ $q_{0.05(2), 27} \approx 3.5$

$LSD = 2.052 \sqrt{19.31 (\frac{1}{10} + \frac{1}{10})} = 4.03$

Tukeys = $3.5 \sqrt{\frac{19.31}{2} (\frac{1}{10} + \frac{1}{10})} = 4.86$

→ A I B
 * MAKE SURE TO DOUBLE CHECK MATH!!!
 Only Advanced and Beginner Different
 ↳ A I B

2/20/2014

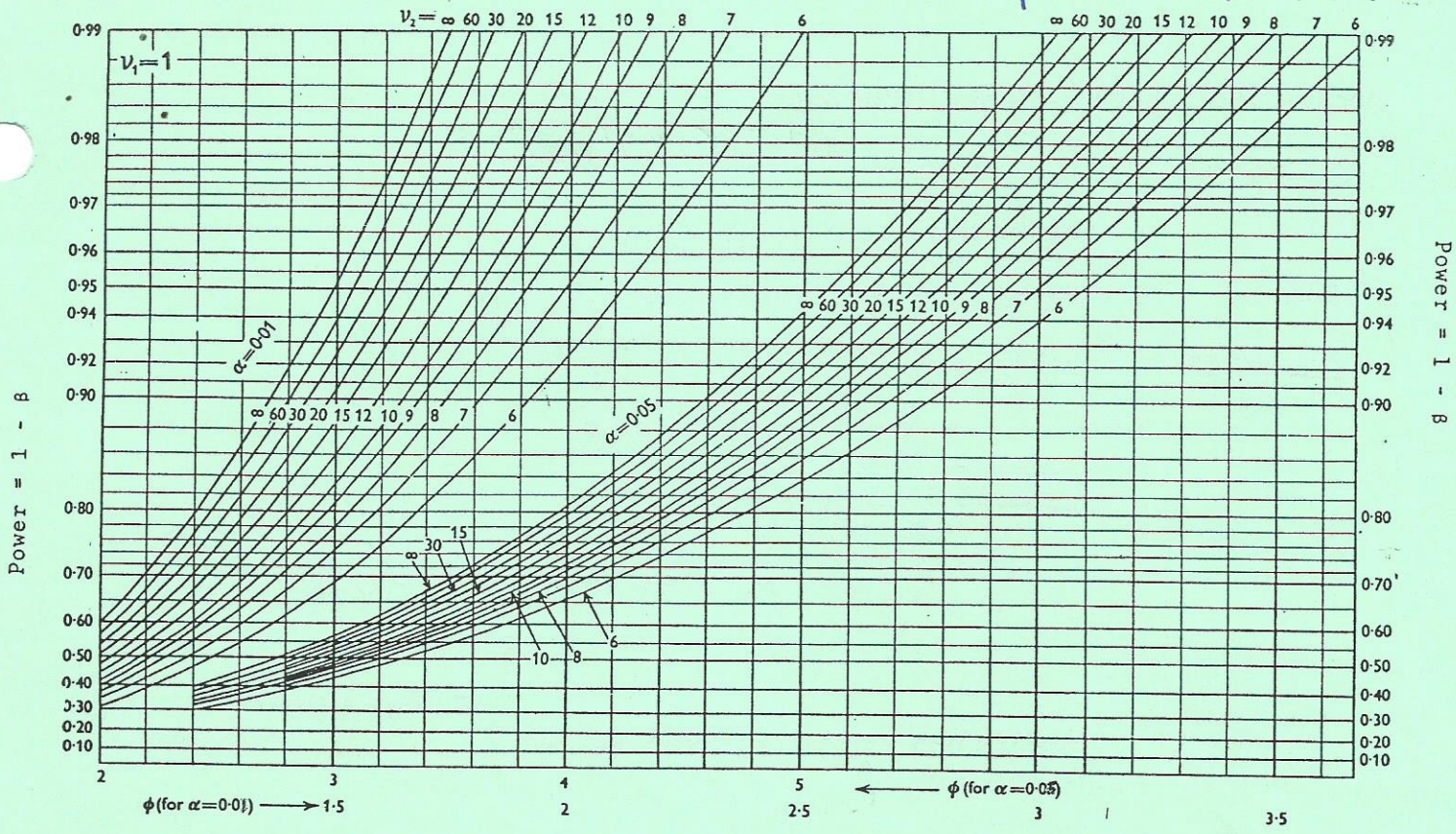


Figure B.1a. Power and sample size in analysis of variance: $v_1 = 1$.

Donuts $v_1 = 3$, $0.05 = \alpha$, $v_2 = 20$ intersect trajectories

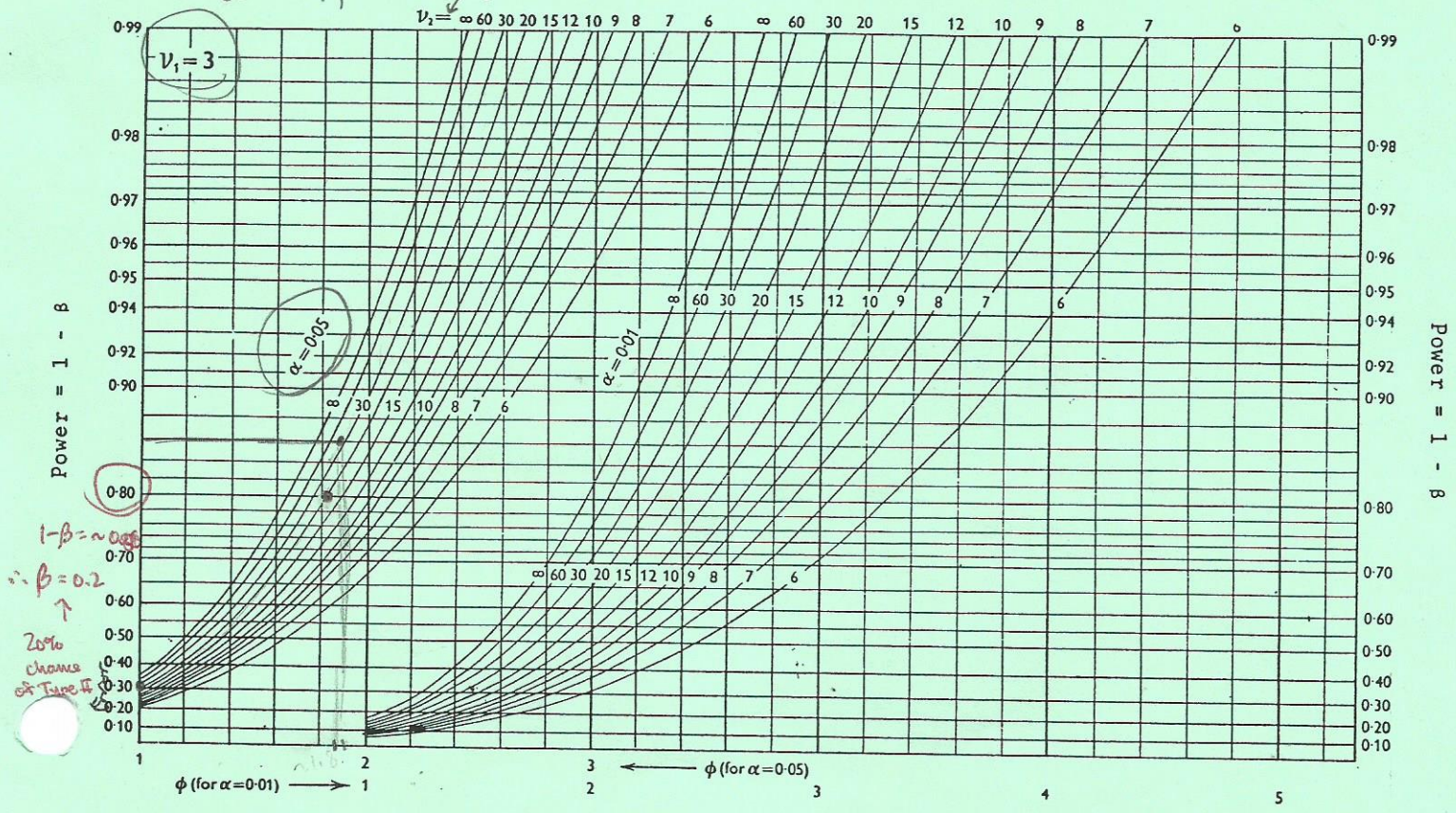


Figure B.1c. Power and sample size in analysis of variance: $v_1 = 3$.

MM

ANOVA and Tukey's HSD of CRUSH data

Response: CRUSH

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
BRAND	3	1796398	598799	5.6106	0.002399
<u>Residuals</u>	44	4696001	106727		Reject Ho

$k = 4$ compute $1 - \beta$ w/ $\alpha = 0.05$

$$\Phi = \sqrt{\frac{(3)(598799 - 106727)}{4(106727)}} = 1.86$$

$$v_1 = 3 \quad v_2 = 44 \rightarrow 1 - \beta = \sim 0.85$$

$\beta = 15\%$ chance
of Type II error

ANOVA has relatively good power

2/20/14

Tukey's HSD (w/ Donuts)

	Lard (85)	Peanut (76)	Corn (72)	Saff (62)	
Lard	-	-	-	-	
Peanut	9	-	-	-	
Corn	13	4	-	-	← Differences b/w means
Saff	23	14	10	-	

$$w = q_{\alpha, k, v_2} \sqrt{\frac{MSW}{2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

↳ ex. $\alpha = 0.05, k = 4, v_2 = 20$

For Donuts HSD

$$w = 3.958 \sqrt{\frac{100.9}{2} \left(\frac{1}{6} + \frac{1}{6} \right)} = 16.8$$

↳ bigger difference required to call two means different

↳ Only Safflower and Lard can be called different by HSD

↳ In essence, we've done a series of q-tests:

$$q_s = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{MSW}{2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim q_{0.05, 4, 20}$$

Ex. Lard vs. Saff. $q_s = \frac{85 - 62}{\sqrt{16.8}} = 5.6$ just for this contrast
 $5.6 > 3.953 \rightarrow$ Reject $H_0: \mu_L = \mu_S$

In literature, we use lines to denote differences in means:

Lard	Peanut	Corn	Saff
85	76	72	62

↳ Means of Lard & Saff different

Many Such MCTs Exist

- Some planned tests - Scheffe's test, Dunnett's Test
- Some unplanned tests - Student-Newman-Keuls test, Bonferroni test
- But know your test before you run it

Estimating Power of a Test

- Statistical testing is based on setting α
- It would be nice to know β (Type II error) in particular if we fail to reject H_0
- If $1 - \beta$ (power) was low then we may want to redo the study or experiment
- Calculating power can be difficult, but "power curves" can help us for ANOVA

Power of a Test: ANOVA

$$\phi = \sqrt{\frac{(k-1)(MS_A - MS_W)}{k(MS_W)}} \quad \text{For Donuts } \phi = \sqrt{\frac{(4-1)(545.5 - 100.9)}{4(100.9)}} = 1.82$$

↳ what do we do w/ ϕ ?

Along w/ $v_1, v_2 \rightarrow$ we can interpolate power using the power curve charts (green)

- For donuts there is 20% chance of getting Type II error (Accept false H_0)

2/20/14

- ↳ With something like this you could actually back calculate specific n for getting a specific $1-\beta$ or determine power for detecting a specific difference b/w means
- Example → Online DSS

2/25/2014

ANOVA Models

ANOVA models are often classified into model types based on how treatments were chosen

The set-up and computations are same but purpose changes

Model I ANOVA (fixed-effects)

treatment levels are deliberately chosen by investigator and are usually fixed and can be manipulated

- You select particular levels for a reason

ANOVA if signif. generally is followed by MCT

Parametric Model:

$$X_{ij} = \mu + T_i + E_{ij}$$

↑ fixed treatment ↑ error

Model II ANOVA (random effects)

treatment levels randomly chosen and cannot be manipulated

- specific treatment levels are less important than just having some different levels

Ex. Choosing trees at random altitudes and measuring growth

A signif. ANOVA suggests that other underlying mechanisms affect response

Usually not followed by MCT; are pretty much just observations

Parametric Model:

$$X_{ij} = \mu + T_i + E_{ij}$$

↑ random effect

Ex. i = random Families trait X_{ij} = antibody level

i	1	2	3	4
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-

$\bar{x}_i = \bar{x}_1 \quad \bar{x}_2 \quad \bar{x}_3 \quad \bar{x}_4$

Do antibodies differ across family?

↳ Don't know why, just want Y/N.

Model III ANOVA (mixed)

Both fixed and random effects; ex. one factor fixed, one random

Requires at least 2 factors in the ANOVA which we will introduce w/ randomized block design

- May combine different experiments and observations

RANDOMIZED BLOCKS ANOVA

2/25/2014

Data: Antibiotic yield (mg) from 4 production processes (treatments), using 5 batches of medium (blocks).

$i =$ Production Process

		A	B	C	D	BLOCK MEAN \bar{x}_j	
$j =$	Medium	1	89 (1)	88 (3)	97 (2)	94 (4)	92
	Batch	2	84 (4)	77 (2)	92 (3)	79 (1)	83
	3	81 (2)	87 (1)	87 (4)	85 (3)	85	
	4	87 (1)	92 (3)	89 (2)	84 (4)	88	
	5	79 (3)	81 (4)	80 (1)	88 (2)	82	
PROCESS MEAN		84	85	89	86	86	

randomization scheme
random 1/4 of medium batch

$H_0: \mu_A = \mu_B = \mu_C = \mu_D$ $H_A: \text{all means not the same}$ $\bar{x} = 86$

ANOVA Table for Randomized Blocks Design

Source	df	SS	MS	F	p-value
Processes	4-1 = 3	$B \sum_{i=1}^B (\bar{x}_i - \bar{x})^2 = 70$	23.33	1.24	~ 0.30
Batches	5-1 = 4	$P \sum_{j=1}^P (\bar{x}_j - \bar{x})^2 = 264$	66.00	3.51	reject H_0 < 0.05
Error	$(4-1)(5-1) = 12$	$\frac{PB}{P} \sum_{i,j} (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2 = 226$	18.83	-	removes 25% ↳ Might not be production process because batches were sign.
Total	19	560	-	-	-

Compare \uparrow

ANOVA Table for Completely Randomized Design

Source	df	SS	MS	F	p-value
Processes	3	70	23.33	0.76	> 0.50
Error	16	490	30.63	-	even lower
Total	19	560	-	-	- blocks help remove variation, increase F-stat ↳ If you can block, do it!

* we want higher F
b/c this means we can correctly reject H_0

For Thursday Class

RBANOVA Example: A researcher studies the effects of four seed treatments (= coatings) and a control (uncoated) group (a total of five treatment levels, T_i) on the germination of soybean seeds. Five germination beds (B_j) were available for the experiment, of little inherent interest to the researcher except that they probably differ from each other in subtle ways. Each replicate is a batch of 100 seeds from which you compute the % germination as the response variable. For this example, set up your experimental design and conduct the analysis of variance, including constructing and interpreting the ANOVA table. You compute the following Sums of Squares ($E = \text{error}$) from the replicates as follows:

$$SS_T = 83.84 \quad SS_B = 49.84 \quad SS_E = 86.56$$

Source of Var	df	SS	MS	F	p
Treatments	$T - 1$	$B \sum_{i=1}^T (\bar{x}_i - \bar{\bar{x}})^2$	SS_T / df_T	MS_T / MSE	
Blocks	$B - 1$	$T \sum_{j=1}^B (\bar{x}_j - \bar{\bar{x}})^2$	SS_B / df_B	$[MS_B / MSE]$	
Error	$(T-1)(B-1)$	$\sum_{i=1}^T \sum_{j=1}^B (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{\bar{x}})^2$	SS_E / df_E		
↓	↓	↓	↓	↓	
T	4	83.84	20.96	3.87	
B	4	49.84	12.46	2.30	
E	16	86.56	5.41		

$F_{0.05, 4, 16} = \sim 3.0$
 $F > 3.00$
 $F_{crit} = \sim 3.0$ **Reject H_0 $p < 0.025$**
 $F < 3.0$ $p > 0.05$

↓
 Block effect was insignificant
 So why block?
 F value for treatment declines

Thought question: You want to test the effect of several different drugs on some human physiological function. Your set of patients (=subjects) vary in age from 20 to 70. You are concerned about how variation in age may affect the response to the drug. What do you do?

2/25/14

Distributing Random Error

- One-way ANOVAs so far have been completely randomized designs
- That is random variation in "subjects" is equally dispersed among the treatments

However, sometimes variation exists that we can't control experimentally

Randomized Blocks ANOVA

- This type of ANOVA removes unwanted variation from uncontrollable sources
 - ↳ remove extraneous variation b/w individuals - like paired-t-test blocks
- Ex. agricultural plots - developed randomized blocks
 - Say you have soil treatments A → E & 20 plots to test fertilizers
 - $n_i = 4$ plots for each fertilizer

Gradient? ↓

A	A	E	C	D
D	C	A	B	A
B	C	D	B	E
E	E	B	C	D

- each plot is a replicate
- randomize treatments (complete)
- potential problem is field is not uniform
 - ex. altitude on each plot, dryness
 - ↳ Something might grow better regardless of fertilizer treatment because of variation

Mechanisms of R-B ANOVA

set-up blocks across some known or suspected gradient

Dry ↓

B	A	D	E	C	Block 1
B	A	C	D	E	Block 2
C	E	D	A	B	Block 3
E	B	A	C	D	Block 4

Wet ↓

Each treatments distributed equally within blocks

Randomize treatments within blocks and use as many blocks as is practical (nT)

Linear Additive Model (for Model III R-B ANOVA)

Parametric Model: $x_{ij} = \mu + \tau_i + \beta_i + \epsilon_{ij}$

↑
random block effect

Estimated Model: $x_{ij} = \bar{x} + (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})$

↑ treatment eff ↑ block effect ↑ block effect error

Partition Variation: $(x_{ij} - \bar{x}) = (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})$

NOTE:

Completely
Randomized
One-way
ANOVA

Treatment Random Err.

$$(x_{ij} - \bar{x}) = (\bar{x}_i - \bar{x}) + (x_{ij} - \bar{x}_i)$$

Treatment Effect τ_i ϵ_{ij}

This simplifies to this term
some error is taken out by the block term
makes F larger, less likely to reject H_0

2/25/2014

Run the ANOVA with fixed treatment effect and random block effect (Model III)

Ex. Let $B = \#$ blocks, $T = \#$ treatments, $MS_W = MSE$

ANOVA Table

Source	df	SS	MS	F	p
Treatments	$T - 1 (v_1)$	$B \sum_{i=1}^T (\bar{x}_i - \bar{\bar{x}})^2$	SST/df_T	MS_T/MS_E	?
Blocks	$B - 1 (v_1)$	$T \sum_{j=1}^B (\bar{x}_j - \bar{\bar{x}})^2$	SS_B/df_B	$[MS_B/MS_E]$	from F stat
Error	$(T-1)(B-1) (v_2)$	$\sum_{i=1}^T \sum_{j=1}^B (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{\bar{x}})^2$	SS_E/df_E		Should compare F-stat here too. Optional - only if you want
Total	$(T \cdot B) - 1$	$\sum_{i=1}^T \sum_{j=1}^B (x_{ij} - \bar{\bar{x}})^2$			to see if variation was indeed pulled out by the block

Assess the two F-values (or at least the fixed F or treatment F) ($p > 0.05 \rightarrow$ no sig var.)

Advantages of R-B ANOVA

- Error SS declines because some of the error is partitioned to blocks
- Generally gives a higher F-stat (= a greater chance of revealing a treatment effect) than complete randomization
- However if blocks remove little or no variation, then R-B can hurt you
- If you include variation, you get less df which means error increases, F plummets, chances of Type II error (β) goes up

2/27/2014

Multi-way Factorial ANOVA

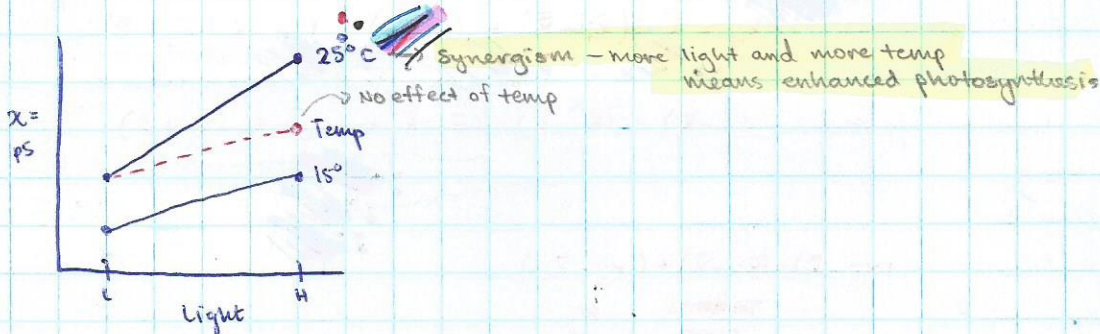
you can simultaneously test for the effects of 2 or more factors with a single multi-way ANOVA

- This is preferable to using several 1-way ANOVAs
 - ↳ Procedure-wise error \rightarrow compounding α
- But you need to consider the effect of one factor on the effect of another factor, called factor interaction

Types of Factor Interaction

Synergism - magnitude of response to one factor increases with increasing value of the other factor

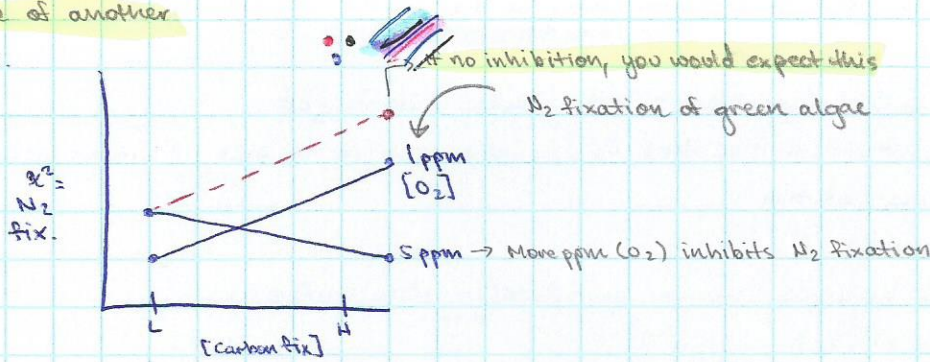
Ex.



2/27/2014

Inhibition - magnitude of response to one factor **decreases** with increasing value of another

Ex.



Interpreting interaction

Parametric Model for Poison-Antidote problem

$$X_{ijl} = \mu + P_i + A_j + I_{ij} + \epsilon_{ijl}$$

poison effect Antidote effect interaction effect Error

Estimated Model

$$X_{ijl} = \bar{x} + (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) + (x_{ijl} - \bar{x}_{ij})$$

Partition Variation

$$(X_{ijl} - \bar{x}) = (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) + (x_{ijl} - \bar{x}_{ij})$$

If you cancel stuff out

$$(x_{ijl} - \bar{x}_i - \bar{x}_j + \bar{x}) \rightarrow \text{error term in}$$

Randomized block ANOVA

By partitioning error into error + interaction, decrease value of error

In the end, decrease MS_E and increase $F \rightarrow$ increase likelihood of rejecting the null hypothesis

ANOVA table for Poison + Antidote problem

R = rep A = antidote # P = poison #

Source	df	SS	MS	F	P
Poisons	P-1	$RA \sum_{i=1}^P (\bar{x}_i - \bar{x})^2$	SS_P / df_P	MS_P / MS_E	? Sig?
Antidotes	A-1	$RP \sum_{j=1}^A (\bar{x}_j - \bar{x})^2$	SS_A / df_A	MS_A / MS_E	? Sig?
Interaction	(P-1)(A-1)	$R \sum_{i=1}^P \sum_{j=1}^A (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2$	SS_I / df_I	MS_I / MS_E	? Sig Interaction?
Error	(PA)(R-1)	$\sum_{i=1}^P \sum_{j=1}^A \sum_{l=1}^R (x_{ijl} - \bar{x}_{ij})^2$	SS_E / df_E	-	-
Total	(PAR) - 1	$\sum_{i=1}^P \sum_{j=1}^A \sum_{l=1}^R (x_{ijl} - \bar{x})^2$			

If there is no significant interaction you can examine the effects of the individual factors (=main effects) and do MCTs

You could take contrast of just poisons or just antidotes

With significant interaction, a specific statement about each factor has limited meaning - you can only say: "The effect of factor A varied depending on Factor B"

Main interaction

Recall that the effect of a factor may be due to the enhancing (or inhibiting) effect of the other factor

- Sometimes such interaction is biologically meaningful
- Be skeptical about studies that fail to interpret (or to present) interaction with the multi-way ANOVA

★ For homework

- MCT for poisons (maybe antidotes) - Tukey's + Fishers

Tukey's for Poisons $k=3$

$$w = q_{\alpha, k, \nu_2} \sqrt{\frac{MS_{sw}}{2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$SS_{sw} = \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \quad MS_{sw} = SS_{sw} / df_{sw} \quad MS_{sw} = 51.7?$$

$\bar{x}_I = 6.2 \quad \bar{x}_{II} = 5.4 \quad \bar{x}_{III} = 2.8 \quad \bar{x}_A = 3.1 \quad \bar{x}_B = 6.8 \quad \bar{x}_C = 3.9 \quad \bar{x}_D = 5.3$
 $SS_{sw} = [(3.1 - 6.2)^2 + (4.5 - 6.2)^2 + (4.6 - 6.2)^2 \dots] + [(3.6 - 5.4)^2 + (2.9 - 5.4)^2 \dots] +$
 $[(2.2 - 2.8)^2 + (2.1 - 2.8)^2 \dots]$
 $= [9.61 + 2.89 + 2.56 + 8.61 + 4 + 23.04 + 6.76 + 1 + 3.61 + 2.89 + 0.01 + 1.96 + 2.89 + 0.81 + 1.6 + 0] +$
 $[3.24 + 1.25 + 1.96 + 9.61 + 14.44 + 0.49 + 0.25 + 49 + 1 + 3.61 + 5.29 + 1.96 + 0.04 + 23.04 + 2.89 + 2.56] +$
 $[0.36 + 0.49 + 1 + 0.25 + 0.04 + 0.81 + 1 + 0.01 + 0.25 + 0.09 + 0.16 + 0.36 + 0.04 + 0.64 + 0.09 + 0.25]$
 $= 132.21$

Use $MS_E! = 2.2$

$\nu_2 = \text{Error df}$

$q_{\alpha, 3, 36} = 3.46$

$w = 3.46 \sqrt{\frac{2.2}{2} \left(\frac{1}{16} + \frac{1}{16} \right)} = 1.3$
 $\bar{x}_I \quad \bar{x}_{II} \quad \bar{x}_{III}$

Apparently I have been doing this wrong.

underline the ones whose means do not differ by Tukey's w

LSD for Poisons $k=3$

$LSD = t_{\alpha, 2, \nu_2} \sqrt{MS_{sw} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$
 $t_{\alpha, 2, 36} = 2.028 \rightarrow LSD = 2.028 \sqrt{2.2 \left(\frac{1}{16} + \frac{1}{16} \right)} = 1.1$
 $\bar{x}_I \quad \bar{x}_{II} \quad \bar{x}_{III}$

Antidotes

$w = q_{\alpha, k, \nu_2} \sqrt{\frac{MS_E}{2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$ now $n=12$

$q_{\alpha, 4, 36} = 3.81 \rightarrow 3.81 \sqrt{0.183} = 1.6$

	\bar{x}_j	A	B	C	D
A (3.1)					
B (6.8)		3.7			
C (3.9)		0.8	2.9		
D (5.3)		2.2	1.5	1.4	

$\square = \text{sign. difference}$

* we could only compare shift separately b/c there was no interaction

A	C	D	B
3.1	3.9	5.8	6.8

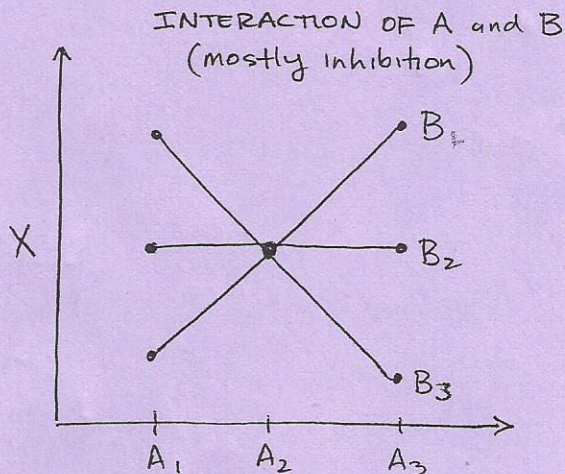
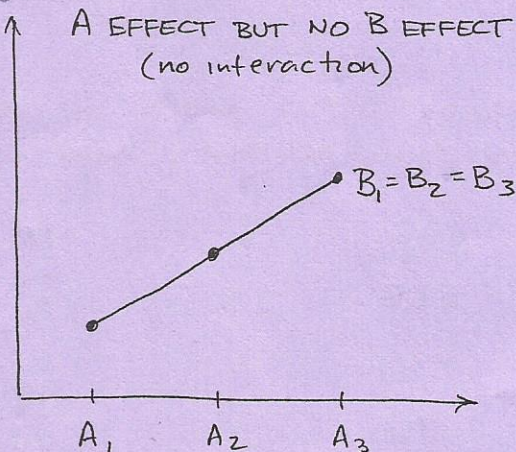
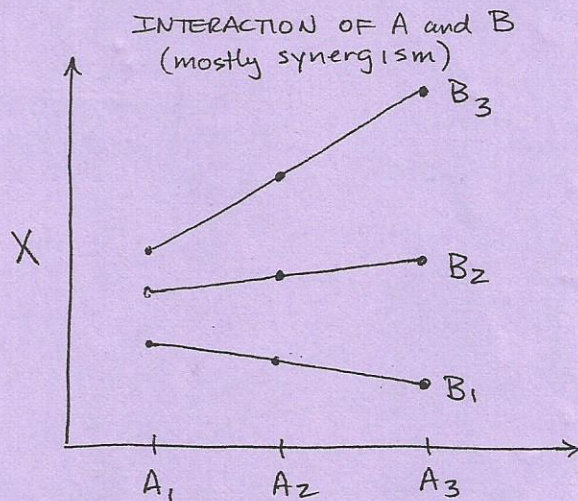
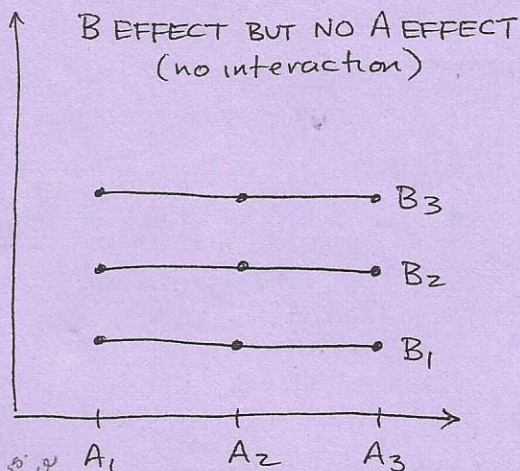
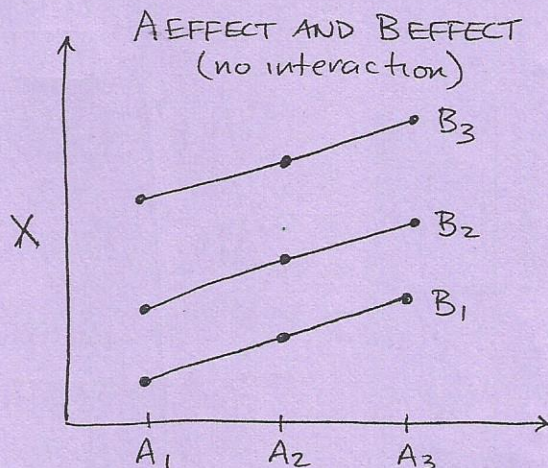
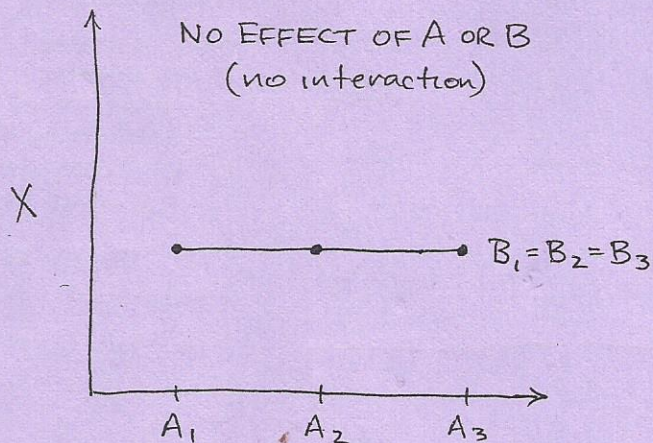
-if there was interaction, best you could

do is compare poison at one level of antidotes (ex. all poisons vs. antidote A)

FACTOR INTERACTION

2/27/14

SIX IDEALIZED ANOVA OUTCOMES FOR 2 FACTORS A and B, each with 3 levels, FOR SOME RESPONSE VARIABLE X.



Does type of M&M have effect on Colors?
 Does type of M&M have effect on price?
 Is there a more economical M&M?
 More mass per package vs. price
 OR
 M in B? S O R G A M
 but time consuming

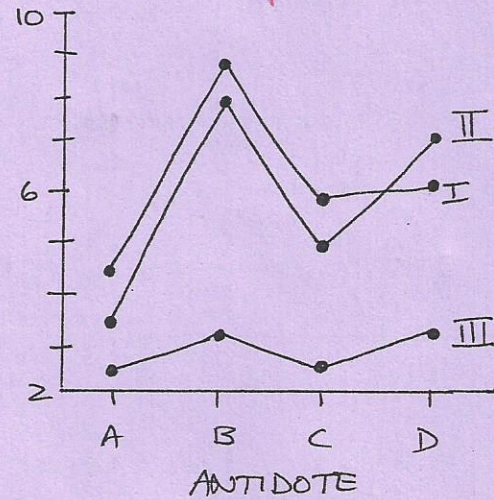
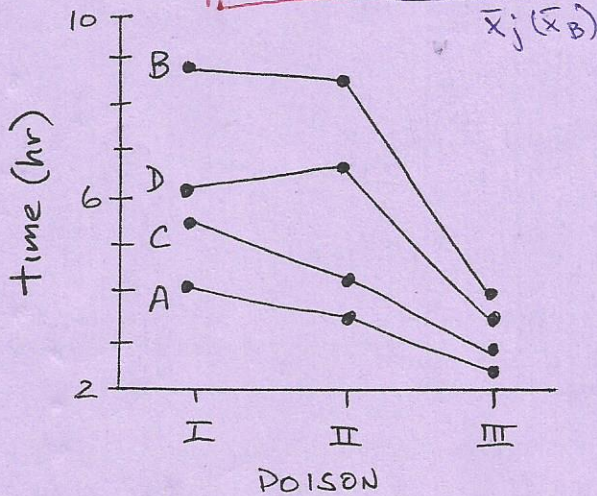
TWO-WAY ANOVA (Multiway or Factorial ANOVA)

Data: Survival times (hours) of insects randomly allocated to 3 poisons and 4 antidotes

$j = \text{ANTIDOTES}$ $l = \text{replicate}$

		A	B	C	D
$i =$ POISON	I	3.1	8.2	4.3	4.5
		4.5	11.0	4.5	7.1
		4.6	8.8	6.3	6.6
		4.3	7.2	7.6	6.2
	II	3.6	9.2	4.4	5.6
		2.9	6.1	3.5	10.2
		4.0	4.9	3.1	7.1
		2.3	12.4	4.0	3.8
	III	2.2	3.0	2.3	3.0
		2.1	3.7	2.5	3.6
		1.8	3.8	2.4	3.1
		2.3	2.9	2.2	3.3

- Features
- fully "crossed"
 - ↳ every factor is combined w/ level of every other factor
 - $k = 12$ cells (3×4)
(12 total treatments)
 - $R = \text{reps per cell} = 4$
 - Balanced
 - $N = 48$ (12×4)



ANOVA Table:

Source	df	SS	MS	F	P
Poisons	2 ν_i	103.3	51.7	23.5	<0.001
Antidotes	3 ν_j	92.2	30.7	13.9	<0.001
Interaction	6 ν_{ij}	25.0	4.2	1.91	ns
Error	36 ν_w	80.1	2.2		
Total	47	300.6			

H_0 : No difference between poisons or antidotes
No interaction between poisons & antidotes

Poison effect sign.
Antidote effect sign.

No interaction b/w P & A

↳ means we can do MCIs of individual poisons or antidotes

3/4/2014

Final Notes: Factorial ANOVA

- Replication w/in cells is required to compute interaction

• • • • • Have to be able to compute cell mean; at least 2 replicates

- Best to have balanced design; cpu can handle missing reps but lose df

- Run MCTs for significant factors

Factorial (Multiway) Experiments

- ANOVA can have any number of factors and any number of levels for each factor; at least 2

• • • • • However the number of measurements needed increases rapidly as design size increases

- Anything beyond 3 factors gets cray-cray

• • • • • An efficient multi-way design is the two level factorial design

where $n = \# \text{ factors}$

	2	3	4	5	6
Main Effect	2	3	4	5	6
# 2-way inter.	1	3	6	10	15
# 3-way inter.		1	4	10	20
# 4-way inter.			1	5	20
# 5-way inter.				1	15
					1

Advantages of 2ⁿ Factorial Design

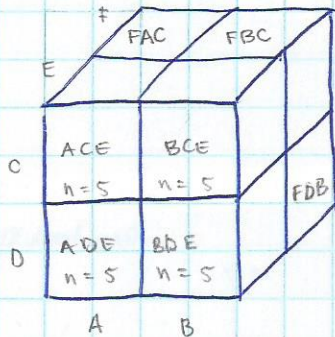
• • • • • Because only 2 levels of each n factor are tested, you need minimum # of measurements to see if a factor has sig. effect

- 2 way interactions are easier to interpret graphically

• • • • • No need for MCT; only 2 levels exist per factor

Factorial Exp. from Yellow handout

$2^3 = 8 \text{ cells}$



- Null hypothesis for plant exp:

Nitrogen, Depth, Temp have no individual effects on decay rate

There is no interaction between the different factors

- Conclusion

Nitrogen and Temp do have individual effect on decay rate

Nitrogen and temp significantly interact

3/6/2014

Major lessons in Experimental Design

- Designing an experiment depends on scientific judgement
- Randomize experimental units among treatments
- Purpose is to eliminate bias, ensure independence

Purposes of Replication

- Estimate Random Error
- Improve precision by reducing the SE
- Increase power by increasing degrees of freedom
- Provide experience with a wider array of experimental units

Assumptions of ANOVA

- Errors are independent (ϵ_{ij})
 - ↳ To ensure this, treatments must be independent of each other
- Errors are about normal in distribution
- Error variance is the same (equal variance) for each cell in the design
- Factor effects are additive, so a linear, additive model describes the data
- REMEMBER L.I.N.E. Linearity, Independent, Normality, Equal variance

Meeting Assumptions

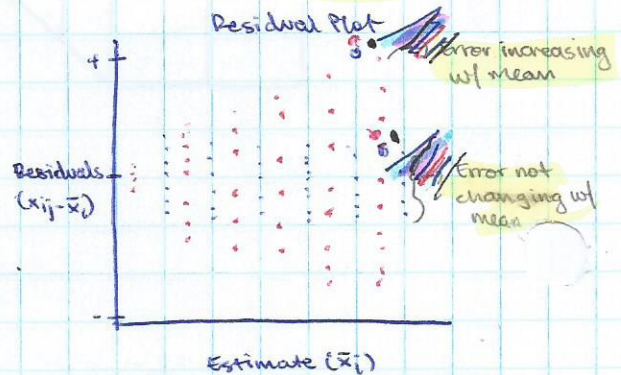
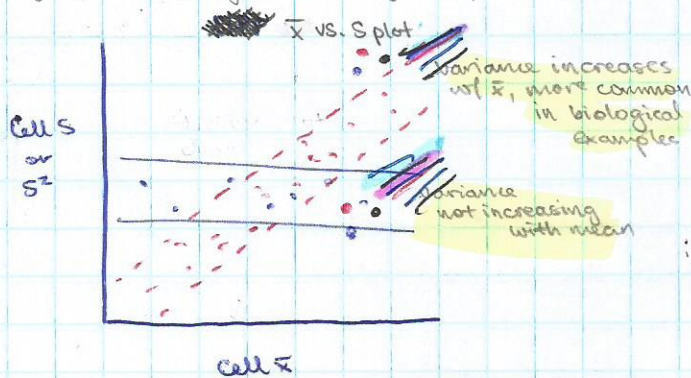
- Assumption 1 (independence): can only be met by randomization before experiment
- Assumptions 2-4: Violation is not as serious as #1, and usually can be "fixed" with a data transformation

Assumption 2: Normality

- Easiest way is to test graphically with normal probability plot (NPP) = plot of normal z scores versus residuals (errors)
- NPP gives straight line if data are normally distributed
- Statistical tests exist to test for normality
 - ↳ Shapiro-Wilk test

Assumption 3: Equal Variance

- Test graphically by examining a plot of a cell means vs. variances (or SD)

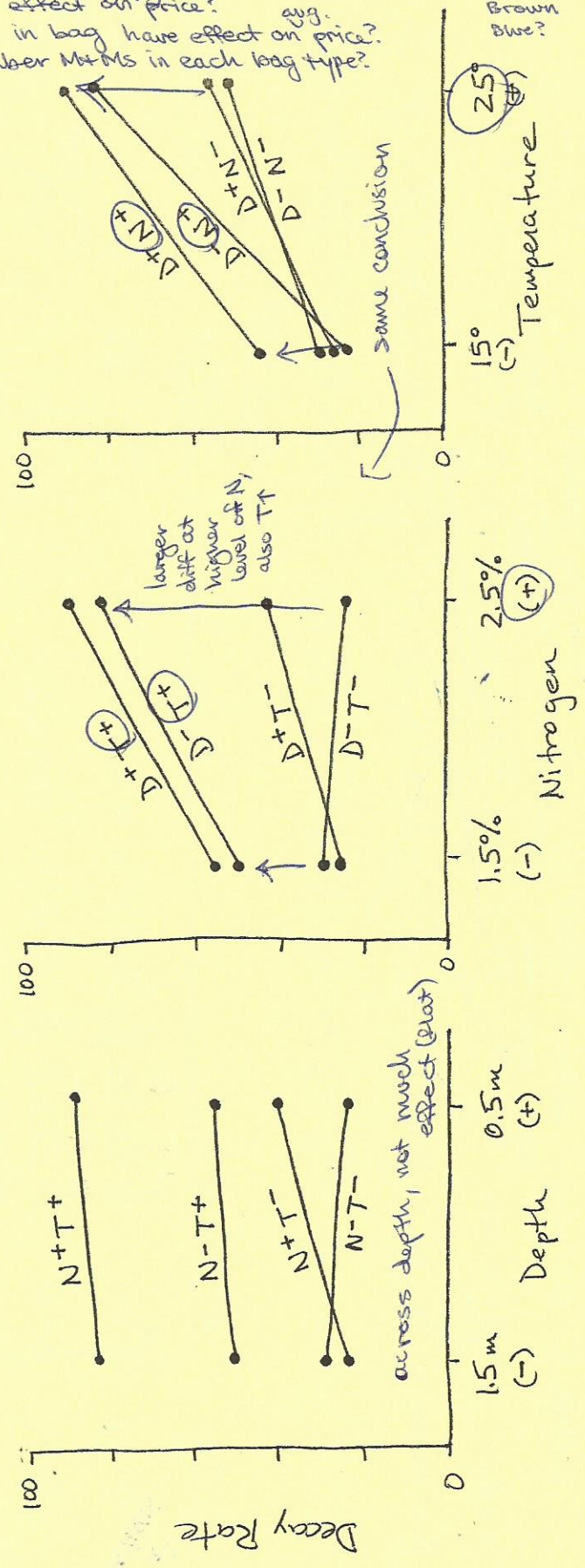


Factors
2³ FACTORIAL EXPERIMENT

Example: Decay rate (parts per thousand and /day) of the aquatic plant Eurasian water milfoil at 2 water depths (0.5 and 1.5 m), 2 water temperatures (15°C and 25°C), and 2 tissue nitrogen concentrations (1.5% and 2.5% of dry weight). For each combination of factor levels, decay was measured for 5 randomly chosen plant samples

D = depth, N = tissue nitrogen, T = temperature
 - and + denote low and high levels of a factor, respectively

Graphical analysis of Interaction:



- For M+Ms
- Does type have effect on color ratio?
 - Does type weight have effect on price?
 - Does number of M+Ms in bag have effect on price?
 - Are there different number M+Ms in each bag type?

3/4/2014

Colors
 Red
 Orange
 Yellow
 Green
 Brown
 Blue?

2³ FACTORIAL ANOVA

Example: decay of water milfoil

Two-way interacts

$$\text{Model: } X_{ijlm} = \mu + D_i + N_j + T_l + I_{ij} + I_{il} + I_{jl} + I_{ijl} + E_{ijlm}$$

depth nitro. temp Random Error
factors replicate

Estimators: $X_{ijlm} = \bar{x} + (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (\bar{x}_l - \bar{x}) + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) + (\bar{x}_{il} - \bar{x}_i - \bar{x}_l + \bar{x}) + (\bar{x}_{jl} - \bar{x}_j - \bar{x}_l + \bar{x}) + (\bar{x}_{ijl} - \bar{x}_{ij} - \bar{x}_{il} - \bar{x}_{jl} + \bar{x}_i + \bar{x}_j + \bar{x}_l + \bar{x}) + (\bar{x}_{ijlm} - \bar{x}_{ijl} - \bar{x}_{ijm} - \bar{x}_{ilm} - \bar{x}_{jlm} - \bar{x}_{ijl} + \bar{x}_{ij} + \bar{x}_i + \bar{x}_j + \bar{x}_l + \bar{x}) + (\bar{x}_{ijlm} - \bar{x}_{ijl} - \bar{x}_{ijm} - \bar{x}_{ilm} - \bar{x}_{jlm} - \bar{x}_{ijl} + \bar{x}_{ij} + \bar{x}_i + \bar{x}_j + \bar{x}_l + \bar{x}) + (\bar{x}_{ijlm} - \bar{x}_{ijl} - \bar{x}_{ijm} - \bar{x}_{ilm} - \bar{x}_{jlm} - \bar{x}_{ijl} + \bar{x}_{ij} + \bar{x}_i + \bar{x}_j + \bar{x}_l + \bar{x}) + (\bar{x}_{ijlm} - \bar{x}_{ijl} - \bar{x}_{ijm} - \bar{x}_{ilm} - \bar{x}_{jlm} - \bar{x}_{ijl} + \bar{x}_{ij} + \bar{x}_i + \bar{x}_j + \bar{x}_l + \bar{x})$

per 2 levels
SS = MS * except

Source	df	SS	F	P
N	2-1=1	NT (2)(5) $\sum_{i=1}^2 (\bar{x}_i - \bar{x})^2 = 313.0$	1.19	NS
T	2-1=1	PT (2)(5) $\sum_{j=1}^2 (\bar{x}_j - \bar{x})^2 = 3547.6$	13.54	sign. N+T effects <0.01
DxN	(2-1) ² =1	DN (2)(5) $\sum_{i=1}^2 \sum_{j=1}^2 (\bar{x}_i - \bar{x})^2 = 15,781.0$	1.06	NS
DxT	(2-1) ² =1	TR (2)(5) $\sum_{i=1}^2 \sum_{j=1}^2 (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2 = 276.7$	0.08	NS
TxN	(2-1) ² =1	NR (2)(5) $\sum_{i=1}^2 \sum_{j=1}^2 (\bar{x}_{il} - \bar{x}_i - \bar{x}_l + \bar{x})^2 = 21.6$	7.18	sign. N+T effects <0.01
DxNxT	(2-1) ³ =1	DR (2)(5) $\sum_{i=1}^2 \sum_{j=1}^2 \sum_{l=1}^2 (\bar{x}_{ijl} - \bar{x}_{ij} - \bar{x}_{il} - \bar{x}_{jl} + \bar{x}_i + \bar{x}_j + \bar{x}_l + \bar{x})^2 = 1882.4$	0.84	NS
Error	2 ³ (5-1)=32	R 5 $\sum_{i=1}^2 \sum_{j=1}^2 \sum_{l=1}^2 \sum_{m=1}^5 (X_{ijlm} - \bar{x}_{ijl})^2 = 8386.6$		
Total	(2 ³ x5)-1=39	$\sum_{i=1}^2 \sum_{j=1}^2 \sum_{l=1}^2 \sum_{m=1}^5 (X_{ijlm} - \bar{x})^2 = 30,430.3$		

*MSE = $\frac{SSE}{dfe} = \frac{8386.6}{32} = 262.1$

sign. inter. also N+T
Can't do MCTB

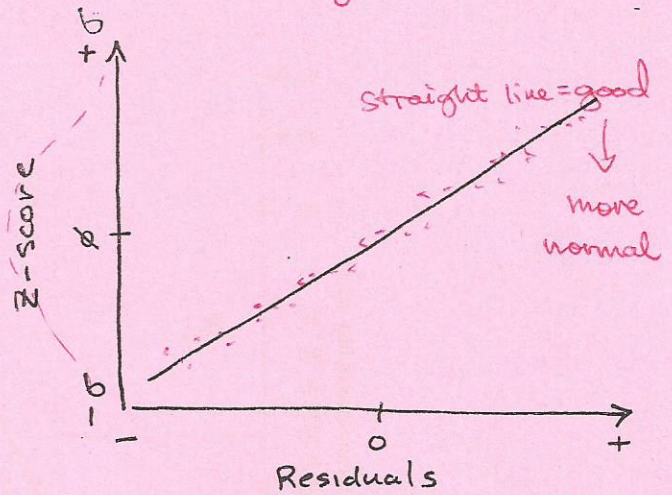
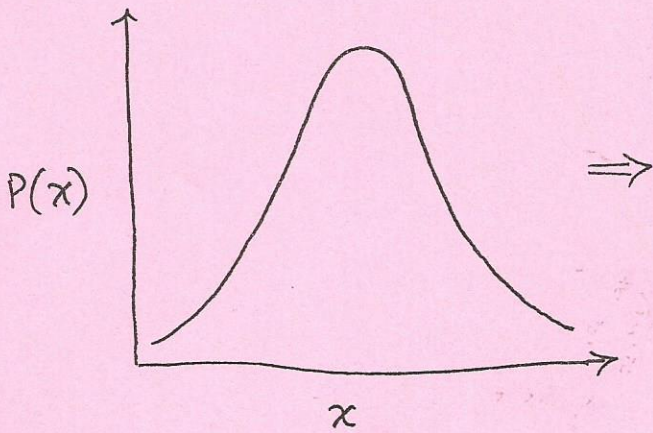
Types of interactions

3/6/2014

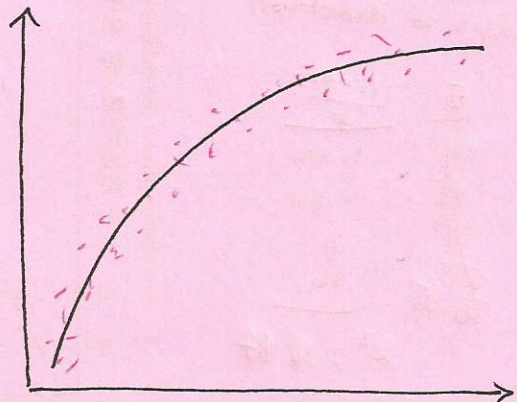
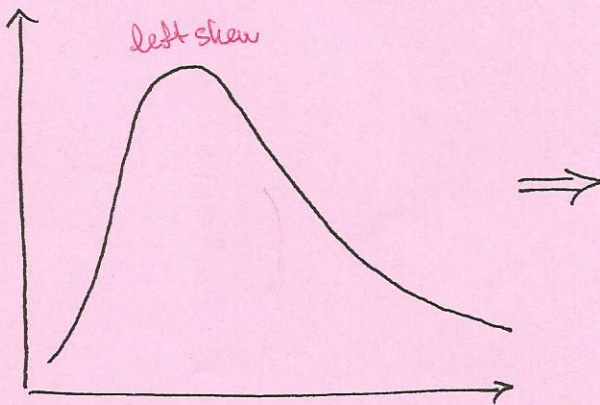
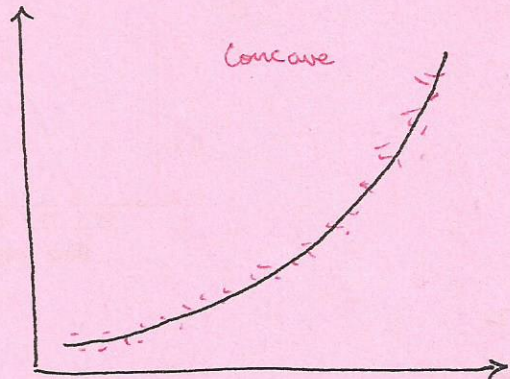
Distribution

Normal Probability Plot

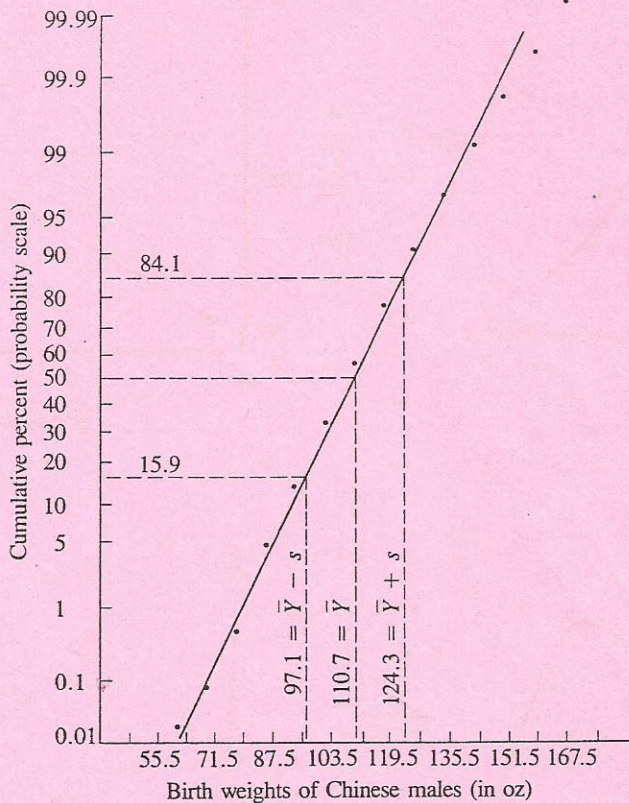
$$Z = \frac{x_i - \mu}{\sigma}$$



$$E = x_i - \bar{x} \text{ or } x_{ij} - \bar{x}_{ij} \text{ etc.}$$



NORMAL PROBABILITY PLOTS



N 8
E 15
W 22
S 36

In Class Problem
to spring break dest.
independent of direction

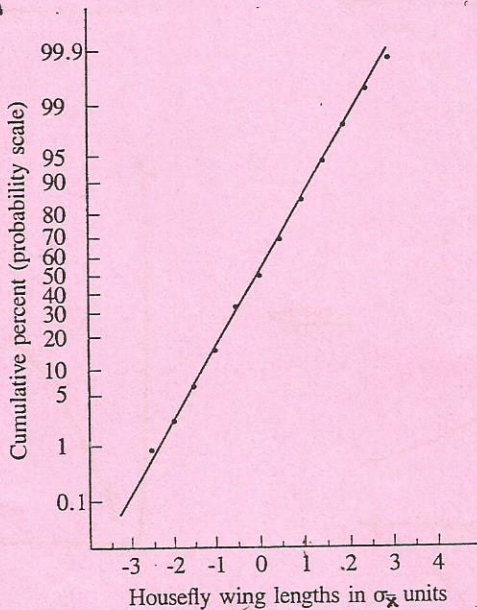
T	E	$\frac{(O-E)^2}{E}$
8	25	11.56
36	25	4.84
15	25	4
22	25	.36
20.15		

$\chi^2 = 20.15$

$df = k - 1 = 4 - 1 = 3$

$\chi^2_{0.001, 3} = 7.815$

$\chi^2 > 7.815 \rightarrow$ Reject H_0 $p < 0.001$



3/6/2014

unequal variances (= heteroscedasticity) violates this assumption

- statistical tests for homogeneity of variances (= homoscedasticity) also exists but they are somewhat weak (ex. Bartlett's Test)

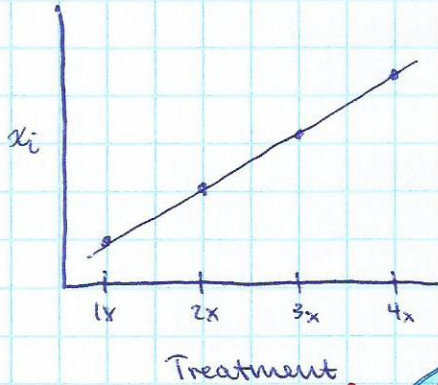
Assumption 4: Linear + Additive

In Multiway ANOVA, significant interaction violates this assumption

↳ This is why it makes it harder to interpret main effects

- on different levels of a factor do not produce a linear response

Ex.



- As we have already learned, evaluation of major effects becomes risky

Options if Assumptions are Violated

- redesign experiment if A#1 violated
- use different statistical test not requiring #2-4 (ex. non-parametric tests)
- Transform the variable to meet assumptions; the same transformation generally satisfies #2-4

Assumption Violated	Consequence	Detection	Rectify
Independence	Bad Experiment	NONE	Re-do experiment
Normality	MS \uparrow , F \downarrow , β \uparrow	Q-Q, NPP, Shapiro-Wilk	Transform Data, NP-test
Equal Variance	MS \uparrow , F \downarrow , β \uparrow	\bar{x} vs. S plot, residual plot	Transform, NP-test
Linear and Additive	Main effects less interpretable	Sign. Inter., or non-linear, non-additive response	Transformation may help; for Multiway ANOVA may need to run 1-way on one factor

- Transformations should be done so normal probability plot fits the shape of the data

↳ ex. for proportions $\rightarrow \arcsin(\sqrt{x})$, for others maybe \sqrt{x}