***INSTRUCTIONS**: Be brief and W R I T E C L E A R L Y . Unless specifically asked for, complete calculations [or even complete sentences] are not needed. Answer in point form when possible. The points for the 14 questions add up to 125. See page 4 re multiple authors. If you work solo, 100%=90 points; if working as a pair, 100%=105 points; if you are a team of three, 100%=115; if a team of four or more, 100%=125. Answers to be handed in at/before the beginning of class on Wednesday May 22*

1 True or false, and explain briefly [2 points each].

a If you add 7 to each value on a list, you add 7 to the SD.

No. Shifting the data by a constant leaves the spread unchanged.

b If you double each value on a list, you double the SD.

Yes. if you multiply inches by 2.54 to get cm, you multiply SD by 2.54

c If you change the sign of each value on a list, you change the sign of the SD.

No. SD is a positive quantity, by definition.

The SD of ages of books is the same as the SD of years of publication, because age = current year (a constant) – year of publication

d If you duplicate each value in a list, you leave the SD approximately unchanged.

Yes. E.g. 80 zeros (0's) and 20 ones (1's) -- SD = 0.4.
160 zeros (0's) and 40 ones (1's) -- SD = 0.4.

e Half the values on a list are always below the mean.

No. it could be 1% or 30% or 70% or 99%.

A very extreme example:

In Canada it might be that

0.990 (99.99%) of people have	*2 legs*
0.008 (0.6%) have	*1 leg*
0.002 (0.25%) have	*0 legs*

ave number of legs
=0.998 x 2 + 0.008 x 1 + 0.002 x 0 = 1.988 legs

Some 99% have an above average number of legs !!!

f In a large set of observations, the distribution of observations follows the Gaussian curve quite closely.

NO. The distribution will be whatever makes sense for the variable in question.

If it is incomes in Canada or the US, they will have a long R tail

if it is lifetimes, they will have a long L tail

and having a bigger sample wont change that!!!

g If two large populations have exactly the same average value of 50 and the same SD of 10, then the percentage of values between 40 and 60 must be exactly the same for both populations.

NO. imagine one has a Gaussian distribution with mean(sd) of 50(10), the other has 1/2 its observations at 40 and the other 1/2 at 60, so it also has means(sd) of 50(10). In this latter distribution, ALL of the values are exactly 1 SD from the mean!

They both have the same means and SD, but very different shapes, and very different percentages of the observations between 40 and 60 (68% in case of Gaussian, 0% in case of 2-point distribution.

See 4 examples in Resources for Ch 1.

h An researcher has a computer file of pre-treatment White blood Counts (WBCs) for patients. They range from 2,800 to 38,600. By accident, the highest WBC gets changed to 386,000. This affects the mean but not the median and the IQR.

Correct. Median is resistant, mean less so.

i The SD of 80 0's and 20 1's is approximately 0.4. The SD of 400000 0's and 100000 1's is also 0.4. [you might use what you said in d]

Correct. Draw the distributions ..

or.. the long way...

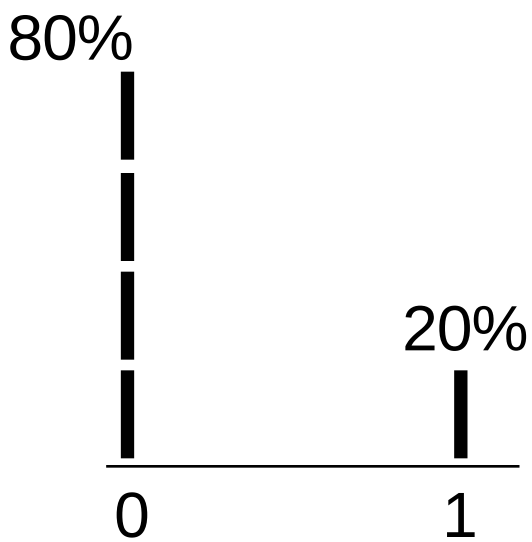
mean = 0.2

80% of squared deviations are $(-0.2)^2 = 0.04$

20% of squared deviations are $(+0.8)^2 = 0.64$

average squared deviation = 0.04 x 0.8 + 0.64 x 0.2 = 0.16

Sqrt(average squared deviation) = 0.4



j A significance test was performed to test the null hypothesis H_0: μ = 2 versus the alternative H_a: μ ≠ 2. The test statistic is z = 1.40. The P-value for this test is thus approximately 0.16

Correct.
2-sided so Prob(Z ≥ 1.4) + Prob(Z ≤ –1.4) = 2 Prob((Z ≤ –1.4) = 2 x 0.0808
{I set it up this way since the Table A gives lower tail area)

k Suppose all students in a class of 20 got the same wrong answer to a multiple choice exam question with 4 choices. To test whether the students colluded [ont triché] while the monitor was out of the room for 2 minutes, the school principal calculated the probability that a random variable Y with a Binomial(20,0.25) distribution would be ≥ 20.

He did this by first calculating μ=20(.25) = 5 and SD= $\sqrt{\frac{0.25 \times 0.75}{20}}$.

He then calculated Prob[$Z \geq \frac{20-\mu}{SD}$] and, finding that the P-value was very small, he concluded that the students had "almost certainly" colluded.

List two problems (1) with the main calculation error and (2) an even bigger logical error in inference; ignore the issue of continuity corrections and the accuracy of the Gaussian approximation.

1. *Principal mixed scales ... used count in the numerator and SD(proportion) in denominator.*

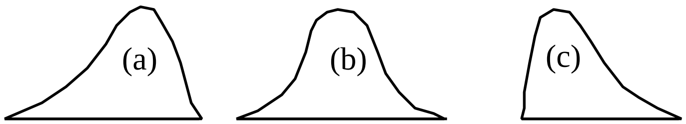
2 *Small P value just says the data are numerically extreme .. not why.*

In the movie (Stand and Deliver, excerpt shown in Against All Odds Video) the teacher had taught the concept incorrectly, and that is why all the students got it wrong.

The phrase "by chance alone" reminds us to ask "what else could it be? Can we rule out all other explanations except "guilty". See also the example of 5 elevated blood sugars in a row.. one might want to check the machine!

2. [5 points]

As part of a survey, a large company asked 1000 of its employees how far they commute to work each day (round trip). The average round trip commute distance was 18 Km, with an SD of 25 Km. Would a rough sketch of the histogram for the data look more like (a) or (b) or (c)? Or is there a mistake somewhere? Explain your answer.



(c). The large SD relative to the mean, the fact that the distances cannot be negative, and that there is no room to have a Gaussian or N(18, 25) distribution without having a large negative component, should point to (c)) as the best answer.

3 [5 points]

In a study of the effects of acid rain, a random sample of 100 trees from a particular forest are examined. Forty percent of these show some signs of damage. Which of the following statements are correct?

(a) 40% is a parameter *NO it is a statistic (from the sample of 100)*

(b) 40% is a statistic *YES*

(c) 40% of all trees in this forest show signs of damage

Can't be sure -- 40% is only an estimate

(d) more than 40% of the trees in this forest show some signs of damage

Can't be sure -- 40% is only an estimate

(e) less than 40% of the trees in this forest show some signs of damage

Can't be sure -- 40% is only an estimate

4 [6 points]

In which of the following would X <u>not</u> have a Binomial distribution? Why?

a. X = number of women in different random samples of size 20 from the 1990 directory of statisticians.

Yes. Binomial(n=20, π = whatever it is)

sampling is without replacement, but 20 is small relative to the size of the directory.

b. X = number of occasions, out of a randomly selected sample of 100 occasions during the year, in which you were indoors. (One might use this design to estimate what proportion of time you spend indoors)

YES.

Binomial (n=100, π = whatever proportion of year is spent indoors)

See the diagram I used in resources in Ch 5.

The sampling calls for independent probes into the Person Time space and so even though there are very clear patterns of inside/outside use, the mechanism by which the sample probes were drawn doesn't use this fact.

So the chance of landing on an indoors person-moment is π for each probe.

c. X = number of months of the year in which it snows in Montreal.

NO

My guess is that the probabilities of some snow in each of the 12 months Jan to Dec are something like this, very disparate, averaging to about 45%.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.98	0.98	0.98	0.90	0.05	0.01	0.00	0.00	0.00	0.10	0.60	0.95

but a binomial with n=12, p = 0.45 would allow some years with few months of snow, and others with many. see Table C with n=12, p = .45;

5 [4 points]

A significance test gives a P-value of .04. From this we can... [indicate True/False for each]

(a) reject H_0 at the $\alpha = .01$ level

No. P-value > pre specified threshold of 0.01

(b) reject H_0 at the $\alpha = .05$ level

Yes. P-value < pre specified threshold of 0.01

(c) say that the probability that H_0 is false is .04. [be careful!]

NO. Not allowed to go from the frequentist

P(data this or more extreme | H_0)

to

P(H_0 | these data)

Frequentists may not speak of a parameter or hypothesis as though it was a random variable with a distribution.

(d) say that the probability that H_0 is true is .04. [be careful!]

NO. for the very same reason.

6 [9 points]

A health department serves 50,000 households. As part of a survey, a simple random sample of 400 of these households are surveyed. The average number of adults in the sample households is 2.35, and the SD is 1.1.

a Sketch a possible frequency distribution showing the variability in the number of adults per household [don't spend a lot of time on trial and error getting the distribution to match the mean and SD exactly; only the general shape is required]

Probably has a long R tail, since the distribution must start at 1, and doesn't leave much room for a symmetric tails on both sides of 2.35. Also know this from experience.

b If possible, find an approximate 95%-confidence interval for the average number of adults in all 50,000 households. If this isn't possible, explain why not.

Sure, ybar (2.35) ± 1.96 SEM

where SEM = 1.1/sqrt(400) ***-- if SRS***

Don't get hung up on fact that y's for individual households may have a skewed distribution. The random variable for inference purposes is the sample mean, not the individual y values. With n=400 the CLT should more than take care of it .. the sampling distribution of all possible sample means should be quite Gaussian.

c All adults in the 400 sample households are interviewed. This makes 940 people. On the average, the sample people watched 4.2 hours of television the Sunday before the survey, and the SD was 2.1 hours. If possible, find an approximate 95%-confidence interval for the average number of hours spent watching television by all adults in the 50,000 households on that Sunday. If this isn't possible, explain why not.

Looks like we could go same route here (even if skewed distribution of no. of hours) BUT the 940 are no longer a SRS -- they are a cluster sample, and the responses from the individuals in same household are likely to be positively correlated. So The SEM is a bit bigger than 2.1/sqrt(940) and would need to be calculated by the appropriate formula for means derived from cluster samples.

7. [6 points]

In a simple random sample (n=225) of all institutions of higher learning in the U.S., the average enrollment was 3,700, with an SD of 6,000. A histogram for the enrollments was plotted and did not follow the normal curve. However, the average enrollment at all institutions in the U.S. was estimated to be around 3,700 (SE = 400).

Say whether each of the following statements are true or false, and explain why.

(a) It is estimated that 95% of the institutions of higher learning in the U.S. enroll between 3,700-800 = 2,900 and 3,700 + 800 = 4,500 students.

NO. 95% CI is for the AVERAGE of ALL institutions. SEM is not for variability of individual institutions.

(b) An approximate 95%-confidence interval for the average enrollment of all institutions runs from 2,900 to 4,500.

YES

(c) If someone takes a simple random sample of 225 institutions and goes two SEs either way from the average enrollment of the 225 sample schools, there is about a 95% chance that this interval will cover the average enrollment of all schools.

YES.. and note the future tense (will cover), invoking the long-run performance of "CI's so constructed"

8 [5 points]

A colony of laboratory mice consisted of several hundred animals. Their average weight was about 30 grams, with an SD of about 5 grams. As part of an experiment, graduate students were instructed to choose 25 animals haphazardly, without any definite method. The average weight of these animals turned out to be around 33 grams. Is choosing animals haphazardly the same as drawing them at random? Discuss briefly, carefully formulating the null and an alternative hypothesis, and computing Z and P.

H0: μ(all ones they might chose) = 30; z = (33-30) / (5/√25) = 3; prob(Z > 3) is only 0.0013; so we have evidence against the claim that they choose randomly. Some reported the small p as alpha < 0.01. Alpha is a fixed cutoff set in advance.

9 [6 points]

An investigator at a large university is interested in the effect of exercise in maintaining mental ability. He decides to study the faculty members aged 40 to 50, looking separately at two groups: the ones who exercise regularly and the ones who don't. There are large numbers in each group, so he takes a simple random sample of 32 from each group, for detailed study. One of the things he does is to administer an IQ test to the sample people, with the following results:

	regular exercise	no regular exercise
sample size	32	32
average score	132	120
SD of scores	16	16

The difference between the averages is "highly statistically significant". The investigator concludes that exercise does indeed help to maintain mental ability among the faculty members aged 40 to 50 at his university.

a State the null and alternative hypotheses, calculate the p-value and verify the statement about the difference being "highly statistically significant"

μ(if exercise) = μ(if do not) vs μ(if exercise) ≠ μ(if do not) [2-sided]

$$t = \frac{132-120}{\sqrt{\frac{16^2}{32}+\frac{16^2}{32}}} = 3$$

Prob($|t_{62df}| > 3$ *) is quite small, around 0.0039, so less than the conventional alpha = 0.05 or 0.01.*

b Is the investigator's conclusion justified? Why/why not?

Not at all; statistically significant just means that unlikely to get this if chance were the ***only*** *factor operating, but given that this is not an experimental study there may be several other factors operating... it may be that those with higher IQ think it helps to exercise. Sig tests are good at ruling out chance but they don't rule in!*
They are concerned with the possible numerical magnitude of chance fluctuations but not with what does cause big fluctuations.

10 [6 points] HIV-1 Transmission

Mastro estimated the probability of HIV-1 transmission, per sexual contact, from female prostitutes to male military prostitutes in Northern Thailand. His conservative estimate of the transmission probability, based on all men, was 0.031 (95% CI 0.025 - 0.040). In a subgroup of men not reporting a history of other sexually transmitted diseases (STDs) his estimate was 0.012 (0.006 - 0.025). He attributes this unexpectedly high value to the possible presence of STDs in female prostitutes (which may have enhanced HIV transmission) and/or high levels of infectivity among the prostitutes who are likely to be at an early stage of HIV infection.

Mastro apparently overlooks these explanations and assumes that the probability of transmission of HIV between regular partners would be the same as that in prostitute-client contacts. He then used this probability of 0.031 to calculate that over 90% of initially uninfected regular partners of seropositive persons would acquire infection over 1 year.

This is inconsistent with data from prospective studies in developing countries suggesting seroconversion rates among HIV- discordant partners of about 10% per year. If it is assumed that couples on average have two sexual contacts per week, then on the basis of simple probability calculations, this gives an average transmission probability per sexual contact of about 0.001 (over 30 times smaller than the conservative estimate of Mastro). [Excerpt from a letter to Editor of The Lancet, May 21, 1994]

a Carry out the calculation using the 0.031 to arrive at an estimate of "over 90%" [paragraph 2]; assume two sexual contacts per week..

104 independent contacts with prob(transmission) = 0.031 per contact
Prob(seroconversion) = 1 – Prob(no conversion in 104 contacts)
= 1 – $(0.969)^{104}$ = 1 - 0.04 = 96% > 90%

b Assuming again two contacts per week, do the reverse calculation that produces an estimate of 'about 0.001' [paragraph 3].

10% seroconversion means 90% remain negative ie Prob(remain neg) = 0.90

Prob(remain neg) = Prob(no conversion in 104 contacts)

= $[\,P(\text{no transmission in 1 contact})\,]^{104}$

= Q^{104} say, { using Q for 1 -prob(transmission in 1 contact) }

Solve $Q^{104} = 0.90$ for Q to give Q = 0.999 [take logs of both sides]

So Prob(transmission in 1 contact) = 1 – Q = 0.001

11 [16 points] TOO TALL?

The US presidential campaign is now gathering pace, with all the customary razzmatazz. Much time and trouble could be saved, however, if the candidates simply had their height measured: with one exception, the President has always been the taller candidate. Tall people are credited with qualities expected of capable people. Nevertheless, all US presidents have been men and almost all studies showing the benefits of tallness have been conducted in male subjects. On casual reading, the paper by Normann and colleagues in Norway on height reduction in tall girls might seem surprising; clearly, positive views about tallness are not shared by Norwegian women. Some girls and their families believe that being "too tall" is a disadvantage, so much so that they are willing to undergo potentially risky treatment.

The Norwegian researchers describe the successful height reduction of 539 girls seen over fifteen years, whose final height prediction was greater than 2.5 standard deviations above the mean (> 181 cm). The population of Norway in 1990 was only 4.24 million, so the large number of patients reported must reflect the tallness of the Scandinavian race (average height 167 cm), the sociocultural definition of normal height, and a general willingness by doctors to treat tallness.... excerpt from Editorial in The Lancet Feb 8, 1992, p339.

Consider the sentences ("The Norwegian researchers... ...to treat tallness") of the second paragraph of the "Too Tall?" editorial above.

a If the relevant population base is 600 ,000 (six hundred thousand) girls, how many of them would meet the inclusion criteria? Use the data given and state any assumptions you make.

2.5 s.d.'s leaves a proportion 0.0062 or 0.62% or 6.2 per 1000 in the upper tail so the estimated number > 2.5 s.d.'s above the mean is 600 x 6.2 or 3720.

Assume Gaussian distribution of heights.

Cannot really assume predicted height = final height; If there are errors in predictions, and if they are random, then SD(predictions) > SD(Final heights)

b From the data given, and any assumptions you make, calculate the 95th percentile of height in the female population.

If 181 is 2.5 s.d.'s or 14 cm above 167, then 1 s.d. must equal 14/2.5 = 5.6 cm,
95th %-ile of Gaussian(0,1) distrn is Z = 1.645, so
95th %-ile of height is 167 + 1.655 x 5.6 = 176.2 cm.

Again, are using assumption of "Gaussian-ness" here (reasonable if Norwegians are ethnically homogeneous) c If the heights of men have a mean of 180 cm, but have the same standard deviation as those of women, what is the probability that a randomly chosen woman is taller than a randomly chosen man?

Let W = Height of woman, M = Height of Man;

Prob (W > M) = Prob (W - M) > 0

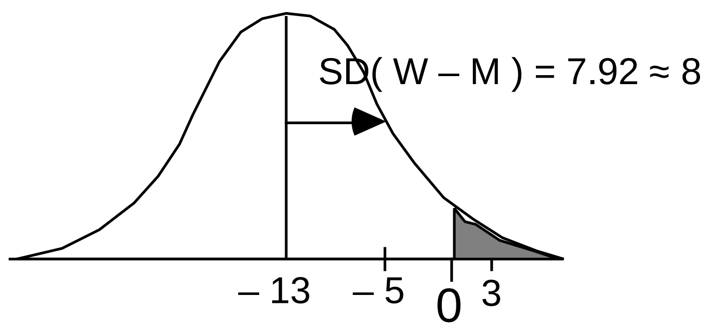
$$E [W - M] = 167 - 180 = -13$$

$$Var [W - M] = 5.62^2 + 5.62^2 = 62.72$$

$$SD [W - M] = \sqrt{62.72} = 7.92$$

$$[= \sqrt{2} \times 5.6]$$

$$z = \frac{0 - \{-13\}}{7.92} = 1.64 ; Prob(Z > 1.64) = 0.0505 \approx 5\%$$



$$z =[0 – \{-13\}] / 7.92 = 1.64$$

Consider the second and third sentences in the first paragraph

d State the hypothesis/claim, implied in the second sentence, in formal statistical terms.

Taller men are winners; shorter men are losers
or
P(Winning | Taller) > P(Winning | Shorter)
{results | determinants}
or
P(Taller | Win) > P(Taller | Lose)
{determinants | results}
or
(NOT SO SHARP): Mean Ht. of Winners > Mean Ht. of Losers

e State the null hypothesis against which you can statistically test the claim.

P(Winning | Taller) = P(Winning | Shorter) = 0.5

or

P(Taller | Win) > P(Taller | Lose)

or

Mean(Height | Winner) = Mean(Height | Loser) i.e. Δ = 0

f There are a few ways to test this; what test statistic you would use?

Compare proportion of Winners in Taller men vs. π = 0.5

or

Compare propn of Taller men among Winners vs. π = 0.5

or

Compare average heights of Winners and Losers

g What reference distribution will your use to describe the sampling variation of the test statistic under the null hypothesis?

Binomial with π = 0.5, n = # of elections
or
t distribution with df = # of elections - 1; pairing avoids the extra variation in heights due to changes in height over 2 centuries.

(Third sentence) The number of elections on which the data are based is not given, but say for the sake of this exercise that it is 50.

h Lay out the steps involved in carrying out the statistical test. You do not need to carry out the detailed calculations but you should give sufficiently clear instructions that a non-statistical assistant could follow them.

- Binomial n = 50 π = 0.5, observe 49/50

  Calculate

  BinProb (49 | π = 0.5) + BinProb(50 | π = 0.5)

  or

  $\text{Prob}(Z \geq \frac{|49 - 25| - 0.5}{\sqrt{50 \times 0.5 \times 0.5}})$ or

  $\text{Prob}((\chi^2_1 \geq \frac{\{|49 - 1| - 1\}^2}{49+1})$

  (the χ^2_1 has easy form when π =0.5 in a "1 x 2" table)

- Paired t–test

  $\text{Prob}(t_{49} \geq \frac{\text{average difference in height - 0}}{\text{SD[differences in height]}/\sqrt{50}})$

- Compare the probability with agreed upon alpha.
  If alt. hypothesis were 2-sided, double the P-value.

12 [10 points]

Refer to the article "Inhibition of oxidation of low-density lipoprotein with red wine" by Kondo K et al. Lancet 344 page 1152 October 22, 1994

a Calculate the SD and variance of the 10 lag times at Day 0 and at Day 14.

SE = SD(indiv)/√10 so SD(indiv)
= SE x √10 = 6.96 and 8.22 respectively. Variances are SD^2

b The authors used 'error bars' of ± **1** SE rather than ± some larger multiple of the SE, presumably so as to ensure that the two CI's for day 0 and day 14 did not overlap. If you were going to put a 95% CI at each point, what multiple would you use?

2.26 because it is based on t_9 *rather than z*

c If the authors calculated a CI for the mean difference as

$$54.7 – 49.1 \pm m \sqrt{2.6^2 + 2.2^2},$$

they would find that for any multiple m bigger than about 1.65, the corresponding confidence interval included a mean difference of zero. Does this mean that the difference is not statistically significant at conventional levels of significance? How does one reconcile this with the p<0.01 reported by the authors?

[Calculating the SE based as $\sqrt{s^2[1/10 + 1/10]}$ where s^2 is the weighted average of the two variances, would give the same SE in this example; so the issue is not one of separate versus pooled variances! Nor is it an issue of 1- versus 2-sided tests!]

They used the paired t-test, which has a much smaller SE (= SD of the 10 differences / √10)

13 [5 points]

Length of published reports

Sir: On completion of a study most researchers want to get their results published at the highest possible level. Do most papers need to be as long as they typically are? Despite editorial attempts to limit article lengths, most authors make statements in the summary, repeat some of them in the introduction, describe them in the materials and results sections, and emphasize and elaborate on them in the discussion. Is all this repetition necessary? Do we as doctors or scientists need to have something repeated 2 or 3 times before we accept or understand it? We think not. We think that papers could be shortened without losing their message if much of the repetition was eliminated, details of standard methods were referenced, and the discussion confined to key elements of the study with a minimum of conjecture.

The Scots have a saying that "guid gear often comes in sma' bulk". Although articles in The Lancet are typically concise (mean 3.9, SD 1.1 columns in November, 1993, vs. 4.2 [1.0] in the British Medical Journal [ns] and 5.7 [1.3] in the New England Journal of Medicine [p < 0.001], one way that more articles, such as short reports, could be published each week would be to ask authors of full papers to reduce them by 300 or 600 words as a condition of publication; they can do this better than most. With more space available for short reports it would become easier to get one published.

Brendan F. Boyce, Brian B. Adamson, Stephen J. Gallacher, James Byars, Stuart H Ralston, Iain T. Boyle Department of Pathology, University of Texas, Health Science Center at San Antonio, San Antonio, Texas, 78284, USA; and Royal Infirmary, Glasgow, UK

From the type of summary statistics he used, assume he did a parametric test. Write out the formula for the test; do you have enough information to carry out it out? If not, what is missing?

2-sample test (for formula see M&M Ch 7) . the 'n' for each sample

14 [20 points]
Refer to the article "Variability of young children's energy intake" by Birch in the NEJM Vol 324: pages 322-325 Jan 24, 1991

a Use hypothetical data to illustrate what the authors mean when they say "the CV makes it possible to compare the variability of measures with different mean values [3rd sentence, statistical analysis portion of methods].

breakfast *200 250 300 mean 250 SD 50 => cv 20%*
lunch *300 400 500 mean 400 SD 50 => cv 13%*

b Explain how the 15 point estimates of Figure 1 were calculated and what they represent.

x axis -- child's body weight

y axis -- average of total intake for 6 days for that child

c Explain carefully how the standard errors in Fig 1 were calculated i.e. what exactly are the components in the SE. Use hypothetical data to illustrate.

SEM = SD(6 daily totals / sqrt[6].

1250 1300 1520 1410 1180 1260 -- SD 124 -> SEM = 124/sqrt[6] = 51

d Why did the authors use SE's in Fig 1? After all, isn't the study about SD's, nor SE's?

If want to show intake vs weight, then the more precisely one can characterize average intake the better .. so SEm good for that

If indeed want to show variability from day to day, then Sd better.

Maybe they think small is better!

e Can you check from the data in the article whether the authors might have mixed up SE and SD?

Could take individual meal data for a subject in Fig 3, reconstruct the 6 daily total and calculate the mean and SD. Then find this child in Fig 1.

f In Fig 2, what does each point represent? Use hypothetical data to show how one lunch data point was calculated.

intakes at 6 lunches

e.g. 250 400 350 300 400 350 -- SD 58 -> SEM = 58/sqrt[6] = 24

g assume that within a particular child (i) across days, the average daily intake is 1500 Kcal (ii) across days, the average intake from each type of meal averages 250 Kcal meal/day (lunch is probably bigger and snacks probably smaller but this simplifying assumption does not change the argument) (iii) for the same type of meal across days, the child's intake from that type of meal has a CV of 30% i.e. a SD of 75 Kcal (iv) the intakes from the six meals in the same or different days are all uncorrelated.

How big would the CV be for the total of the 6 intakes in a day?

meal: CV = 30% so SD itself = 30% of 250 = 75.

SD[total]= sqrt[sum of 5 variances for indiv. meals]

= sqrt[75^2 + 75^2 + 75^2 ... + 75^2] = 183

ave[total] = 1500, so CV[total] = 183 / 1500 = 12.2% (= 30%/sqrt[6] !)

Compare your answer with the reported mean within subject CV for total daily energy intake [10.4%, end of 2nd paragraph of results].

The 12% is not THAT much bigger than what you would get with no correlation from meal to meal.

Comment on whether the assumption of strong negative correlation between energy intake at one meal and the next meal [see next paragraph] is needed to explain the observed CV for total daily intake.

Maybe a little bit, but already down to 12% without it

h Write a short statistical paragraph to the editor explaining how the smaller CV for total daily intake that the individual meal intake could be expected, even if there were no negative correlation. Try to use words rather than formulae.

Most of the seeming "tight" day to day variation in total intake could be explained by the sqrt[6], the same factor that tightens sums and averages!!