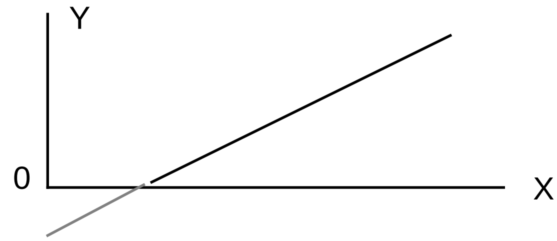
2.1 (a) 95%CI for β_1 excludes zero; that is same as saying that a test of $\beta_1=0$ carried out with $\alpha = 0.05$ (using two sided P-value) would be statistically significant. So, yes, conclusion is correct. Fact that 95% CI excludes zero implies $P < 0.05$.

(b) Indeed, sales cannot be negative, and so mean sales when X=Population=0 cannot be negative. BUT, this may simply be an artifact of where the data are i.e. if the data points are far away from X=0, then any estimate of the intercept will be very imprecise. Remember the formula for the variance of b_0 -- it has a term in it that is proportional to the square of xbar! It could also be negative if the true relationship were



2.2 No. H_0 includes <u>negative</u> slopes.

2.4 (a) CI for β_1 excludes zero; indeed it suggests that β_1 is positive. If β_1 were zero, there wouldn't be much point in having entrance tests!

(b) Can use test and 100(1-α)% CI and 2-sided test with size α), interchangeably. They both involve the same standard error and the same critical value of the reference distribution.

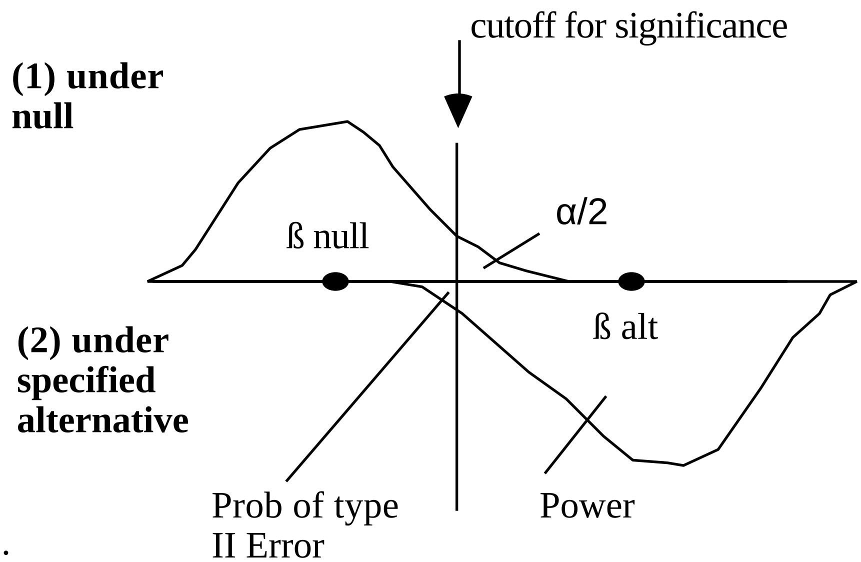
(c) ALWAYS state whether the P-value is 1- or 2-sided!

2.6 (a) Some of you have problems with the interpretation of a CI. Some of you use the idea of (future) repeated sampling and up to a point that is ok, but it would not be much re-assurance for a client who came to you once to hear you speak of repeating the procedure in the future. Also, remember that the client is not interested in the (future or otherwise) behaviour of b_1 per se, but rather (via b_1) in β_1 ! I know that in order to make β_1 the <u>subject</u> of the sentence, you would need to approach the problem from a Bayesian perspective. You can partly get around these issues by speaking about the performance of the "CI -setting procedure", i.e. by saying that 95% of "CI's so constructed" do well. Better not to talk about the future, but about the past (and present) performance of CI's in the hands of "CI-makers".

A few of you mis-spoke quite badly when you wrote something " The CI is x.x to y.y; if repeated samples were taken, and CI's calculated, in 95% of them the true β_1 would be would between x.x and y.y". Can you see why this is **quite inaccurate**?

(d) H_0: $\beta_1 \leq 9$ versus H_a: $\beta_1 > 9$ ... so a **<u>1</u>**-sided test. But note that in (c) you were asked for a (2-sided) 95%CI; if the lower limit of this exceeded 9, then this would be same as finding that H_0 is rejected in a **1**-sided test with α=0.025. So, can use interchangeability of tests and CI's.

(e) Power: Here there are two key steps (1) establish what the b_1 cutoff is for a test to be significant (2) calculate where this cutoff is with respect to the distribution of b1 under the alternative. I find myself making the following drawing for such calculations:



:

Some of you forgot that the null value in part (d) was β_0 = **9**, and not β_0 = 0.

2.9 The computer program cannot anticipate which value of X the user is interested in. Most packages will calculate and show the intervals corresponding to the values of X in the dataset, and so if the X_h of interest is near or between some of these X values, you could extra(inter)polate.

2.10 A few of you thought that (a) implied the <u>mean</u> level when T=31. But the question was about a specific tomorrow. One could imagine many different days when one could set T to 31, and what the <u>mean</u> of the Y's for <u>all</u> of these days would be. The "tomorrow" in this example is just <u>one</u> of the possibilities. There was no dis-agreement that the object in (b) was the mean. For (c), I took it to mean one specific month, not the average for several such months.

2.11 Think of mean response i.e. $E[Y \mid X_h]$ as the case where m is infinite.

The variance has the form $\sigma^2\{1/m + 1/n + (X_h - \bar{x})^2 / \Sigma(x - \bar{x})^2\}$. Our prediction intervals were for cases of m=1, whereas our CI's were for the case of m = ∞. This question is for the intermediate case, e.g. asking what would be the mean weight of say a team of m=11 individuals.

2.12 No; individual responses maintain their individuality and non-predictability; statistics (numbers based on aggregates) can indeed be made arbitrarily precise by increasing n. It is a little like the difference between predicting the temperature for September 30. Even if we have data on the last 100 September 30's, that just means that we have a precise estimate of the expectation, and we cannot reduce the variability of the 100 individual days (maybe there is a small temporal trend, but it is tiny compared to what we don't know!)

You all saw that $\sigma^2\{1/n\}$ goes to 0 as n –> ∞ but some of you missed the

fact that $\sigma^2 (X_h - \bar{x})^2 / \Sigma(x - \bar{x})^2$ does as well, You can see this better if you write

$$(X_h - \bar{x})^2 / \Sigma(x - \bar{x})^2$$

as

$$(X_h - \bar{x})^2 / \{ n \times \text{var(x's in sample)} \}$$

which clearly goes to 0 as n –> ∞.

2.13 (a) point est. of mean Y at X=47 : -1.7 + 0.84(4.7) = 2.25

estimated var(...) $= \text{MSE}\{ 1/20 + (4.7- 5.0)^2 / [19 \times \text{SD}[20 \text{ x's}]^2]$

$= 0.189\{ 1/20 + (-0.3)^2 / [19 \times 0.69^2] \} = 0.0113$

i.e. SE(...) = 0.11

95% CI = 2.25 ± $t_{18,95}$ SE = 2.25 ± 2.101(0.11) = 2.25 ± 0.22

(b) point est. of individual Y at X=47 : -1.7 + 0.84(4.7) = 2.25

estimated var(...) $= \text{MSE}\{1+1/20 + (4.7- 5.0)^2 / [19 \times \text{SD}[20 \text{ x's}]^2]$

$= 0.189\{ 1 + 1/20 + (-0.3)^2 / [19 \times 0.69^2] \}$

= 0.1946 ***(most of variance is from individual)***

i.e. SE(...) = 0.44

95% interval = 2.25 ± $t_{18,95}$ SE = 2.25 ± 2.101(0.44) = 2.25 ± 0.92

(c) same issue as 2.12 above.

NOTE: Instead of using the variance formula

$$\mathrm{var}(\hat{Y}) = \sigma^2\{1/n + (X_h - \bar{x})^2 / \Sigma(x - \bar{x})^2\},$$

a more useful form, used by some of you, is

$$\mathrm{var}(\hat{Y}) = \sigma^2/n + (X_h - \bar{x})^2 \sigma^2 / \Sigma(x - \bar{x})^2$$

$$= \mathrm{var}(\bar{y}) + (X_h - \bar{x})^2 \mathrm{var}(b_1)$$

This come from the representation:

$$\hat{Y} = \bar{y} + (X_h - \bar{x})\ b_1$$

and the fact that $\bar{y}$ and b_1 are statistically independent.
.

(d) See the comment at end of 1st paragraph of page 68. Unfortunately, most packages just give the CI for the E[Y| X] at each separate X as a band. But this is not really the band for the line. The CI's for each separate X value are not statistically independent -- one can be 95% confident about any the CI for any <u>one</u> specified X only. Thus the procedure for the <u>entire</u> line. Because it is completely general, and assumes that the entire line from x = -∞ to X = +∞ is of interest, it is somewhat conservative.

2/3 Analysis of Rates of Fatal Crashes

	55 mph					65 mph
YEAR	1982	1983	1984	1985	1986	1987
Rate per10^8 vehicle miles	2.8	2.0	2.1	1.7	1.9	2.9
N OF YEARS	5					1
MEAN(Rate)	2.100					2.9
VARIANCE(Rate)	0.175					0.0

DEP VAR: Rate N:5 MULTIPLE R:0.794 MULTIPLE R^2: 0.630

STANDARD ERROR OF ESTIMATE: **0.294**
(This "STANDARD ERROR OF ESTIMATE" is a misnomer; It is the square root of the average squared residual and might be called the "average residual")

VARIABLE	COEFF.	STD ERROR	T	P(2 TAIL)
CONSTANT	**418.740**	**184.345**	2.272	0.108
YEAR	**-0.210**	**0.093**	-2.260	0.109

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	0.441	1	0.441	5.108	0.109
RESIDUAL	0.259	3	0.086		

---

a Interpret the fitted "constant" of 418.740.

The fitted line extrapolated back to the year A.D. 0 !

Why does it have such a large standard error?

Because a small error in estimating the slope is propagated when multiplied by 1980 or so (remember thatwe do not get to the present data by starting at A.D. 0 and projecting forward 1980 years; rather the b_0 is a backwards extrapolation!

Rewrite the fitted model using a more appropriate "beginning of time" (don't worry about being Y2K compliant! you could even use the Microsoft definition of the "beginning of time").

We could choose our origin as 1982 or some such year. Fitted equation is then
2.52 - 0.210 (year - 1982)

b Interpret the -0.210 and its standard error 0.093 [for parts a and b use your parents in law as your intended readership]

difference in fitted means 1 year apart. SE is a measure of precision, reflecting the small sample size, but also the small residuals.

c Scientists often interpret an absolute value of "b / SE(b)" of 2.0 or more as "P<0.05(2-sided)". Here b/SE(b) is -2.26, but P(2 tail) is 0.109!! Explain.

Only 3 df so instead of $Z_{95} = 1.96$, have $t_{95} = 3.18$.

d Use equations 2.4 and 2.4a (p46) to quickly hand-calculate the b_1. What weights do the 5 different rates receive in the calculation? Why are these weights appropriate?

weights for 5 y's are -2/10, -1/10, 0, 1/10 and 2/10 reflecting the greater value of the more extreme X's

e Obtain the 5 fitted values and thus verify by hand that the 0.294 is in fact the square root of the "average" squared residual.

Fill in the blanks... *2.52 - 0.21(1987 - 1982) = 1.47;*

Interval(centered at 1.47)

$= RMSE \times t_{3,95} \times sqrt[1 + 1/5 + 9/10] = 0.294 \times 3.18 \times 1.45 = 1.355$

A number of you (thinking 1-sided ??) used the $t_{3,90}$ rather than $t_{3,95}$

4 Alcohol and impairment

a Since we have the means and SD's for before and after, we could perform an ***unpaired*** t-test

$$t_{22} = \frac{\bar{y}_{before} - \bar{y}_{after}}{\sqrt{s^2_{before} / 12 + s^2_{after} / 12}} = \frac{58.5 - 43.0}{\sqrt{11^2 / 12 + 13.6^2 / 12}}$$

However, this test is insensitive, as the SE in the denominator contains between-person variances in their responses before and after [as if we studied 12 persons before and a ***different 12 after***] . We should focus on the 12 within-person ***changes*** i.e. on the 12 differences

$$d = y_{before} - y_{after}$$

yielding the paired t-test statistic

$$t_{11} = \frac{\bar{d}}{\sqrt{s^2_d / 12}} = \frac{58.5 - 43.0}{\sqrt{??^2 / 12}}$$

(The non-parametric analogues such as signed rank test or straight sign test are also appropriate).

The numerator of the test is the same, since the average difference equals the difference of the averages; but the denominator, involving

$$var[y_{before} - y_{after}] = var[y_{before}] + var[y_{after}] - 2\, covar[y_{before}, y_{after}]$$

is likely to be smaller, since y_{before} and y_{after} are likely to be positively correlated.

We do not have the 12 absolute changes d_1 to d_{12} but we do have the % change in each person, i.e., we have

$$\%d = 100\{ y_{before} - y_{after} \} / y_{before}$$

for each subject, so we can do the paired test on these instead.

b Can use $t=r[(n-2)^{1/2}] / [(1-r^2)^{1/2}]$. Note that this just tests the H_0: $\rho = 0$. Note that this is the same test as the test of $\beta_1 = 0$

c I would say nonzero intercept because (i) might have threshold effect (ii) experiment did not measure baseline alcohol (which might not be exactly zero) (iii) could be learning or tiring effect.

d Point to ss #3,5 and 6. Point out that the line is an estimated line, with a lot of uncertainty (different in the next 12 ss) and that ***even then*** it refers only to the average person; one must still allow for individual variation. Even if we had data on 120 or 1200 persons, they would only allow us to estimate the mean precisely... they still would not allow us to predict precisely how the next person (you or I) would respond.

Remember the question (elsewhere in homework) about being able, by increasing n, to ***reduce the uncertainty concerning the mean, but not the individual.***

e Measure x and y before and during (that way could get an estimate of "y at 80" for each person. After all, that is what the main question was. Fit a separate curve for each person and then describe the distribution of "%impairment at 80" estimates.

Measure personal characteristics and see if the estimates of impairment segregate along these lines.

Measure same persons in control situation (non alcohol) to understand how much could be simply due to tiredness etc.