

## Composite scores

### 1. The Mean of Composite Scores

A composite score is the score which results from summing two or more scores. Composite scores will be symbolised as  $C$ .

It can be shown that the mean of a composite equals the sum of the means of the components. Using  $\bar{C}$  as the mean of the composite and

$\bar{X}_1, \bar{X}_2$ , etc. for the means of the components we have:

$$\bar{C} = \bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_n \quad (9:1)$$

*Proof*

$$(1) \quad \bar{C} = \frac{\sum C}{N}$$

$$(2) \quad \sum C = \sum (X_1 + X_2 + \dots + X_n)$$

(3) By Summation Rule 1. (2) is equal to

$$\sum X_1 + \sum X_2 + \dots + \sum X_n.$$

(4) Therefore:

$$\frac{\sum C}{N} = \frac{\sum X_1}{N} + \frac{\sum X_2}{N} + \dots + \frac{\sum X_n}{N}$$

(5) The values on the right are of course means so

$$\bar{C} = \bar{X}_1 + \bar{X}_2 + \dots \bar{X}_n$$

$S_{xy}$

Sometimes a composite score is the sum of weighted components. For example  $C$  might equal  $2X_1 + 1.5X_2 + kX_3$ , the weights being 2 for  $X_1$ , 1.5 for  $X_2$ , and  $k$  for  $X_3$ . Let us symbolise weights as  $W_1, W_2 \dots W_n$ , then:

$$\bar{C} = W_1\bar{X}_1 + W_2\bar{X}_2 + \dots W_n\bar{X}_n \quad (9:2)$$

*Proof*

$$(1) \quad C = (W_1X_1 + W_2X_2 + \dots W_nX_n)$$

$$(2) \quad \sum C = (\text{using Summation Rule 1})$$

$$\sum W_1X_1 + W_2X_2 + \dots \sum W_nX_n$$

$$(3) \quad \text{So } \frac{\sum C}{N} = \frac{\sum W_1X_1}{N} + \frac{\sum W_2X_2}{N} + \dots \frac{\sum W_nX_n}{N}$$

$$(4) \quad \text{So } \bar{C} = W_1\bar{X}_1 + W_2\bar{X}_2 + \dots W_n\bar{X}_n$$

## 2. The Covariance

Before considering the variance of a composite it will be worth recalling the covariance and some computational formulae connected with it. The covariance has been mentioned before in the chapter on correlation. It is defined as the mean of the products of subjects' deviation scores on two tests. Using  $S_{xy}$  as the symbol for the covariance:

$$S_{xy} = \frac{\sum(X - M_x)(Y - M_y)}{N} = \frac{\sum xy}{N} \quad (9:3)$$

It will be recalled that one formula for  $r_{xy}$  was:

$$r_{xy} = \frac{\sum xy}{N\sigma_x\sigma_y} \quad (9:4)$$

From this it follows that:

$$S_{xy} = \frac{1}{N} \sum xy = r_{xy}\sigma_x\sigma_y$$

That is the mean product of the deviation scores equals the product of the correlation coefficient and the two standard deviations. A convenient formula for the covariance is:

$$S_{xy} = \frac{\sum XY}{N} - M_x M_y \quad (9:6)$$

*Proof*

$$(1) \quad (X - M_x)(Y - M_y) = XY + M_x M_y - M_x Y - M_y X$$

$$(2) \quad \sum(X - M_x)(Y - M_y) = \sum XY + NM + NM_x M_y - M_x \sum Y - M_y \sum X$$

(3) Dividing by  $N$  gives:

$$\frac{\sum (X - M_x)(Y - M_y)}{N} = \frac{\sum XY}{N} + M_x M_y - M_x \frac{\sum Y}{N} - M_y \frac{\sum X}{N}$$

(4) But  $\frac{\sum Y}{N} = M_y$ ; and  $\frac{\sum X}{N} = M_x$

So the right hand term becomes:  $\frac{\sum XY}{N} - M_x M_y$

### 3. The Variance of a Composite Score

By now the formula for the variance is familiar, i.e.

$$\frac{\sum (X - M_x)^2}{N}$$

so the formula for the variance of a composite will be:

$$\begin{aligned} \sigma_c^2 &= \frac{\sum (C - \bar{C})^2}{N} \\ &= \frac{\sum [(X_1 + X_2 + \dots + X_n) - (\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_n)]^2}{N} \end{aligned} \quad (9:7)$$

The last term can be written in deviation scores:

$$\sigma_c^2 = \frac{\sum [(X_1 + X_2 + \dots + X_n) - (\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_n)]^2}{N} \quad (9:8)$$

$$= \frac{\sum [(X_1 - \bar{X}_1) + (X_2 - \bar{X}_2) + \dots + (X_n - \bar{X}_n)]^2}{N}$$

$$= \frac{\sum (x_1 + x_2 + \dots + x_n)^2}{N}$$

An easy way to work out all of the values involved in  $(x_1 + x_2 + \dots + x_n)^2$  is to prepare a square table thus:

	$x_1$	$x_2$	$x_3$	...	$x_n$
$x_1$					
$x_2$					
$x_3$					
...					
$x_n$					

The body of the table is formed by multiplying the marginal elements. As follows:-

	$x_1$	$x_2$	$x_3$	...	$x_n$
$x_1$	$x_1^2$	$x_1x_2$	$x_1x_3$	...	$x_1x_n$
$x_2$	$x_1x_2$	$x_2^2$	$x_2x_3$	...	$x_2x_n$
$x_3$	$x_1x_3$	$x_2x_3$	$x_3^2$	...	$x_3x_n$
...	...	...	...	...	...
$x_n$	$x_1x_n$	$x_2x_n$	$x_3x_n$	...	$x_n^2$

Thus for each individual we have:

(a)  $x_1^2 + x_2^2 + \dots + x_n^2$  and

$$(b) \quad 2x_1x_2 + 2x_1x_3 + \dots 2x_{(n-1)}x_n$$

Summing across individuals and dividing by  $N$  gives:

$$\begin{aligned} & \frac{\sum (x_1 + x_2 \dots x_n)^2}{N} \\ &= \frac{\sum (x_1^2 + x_2^2 + \dots x_n^2 + 2x_1x_2 + 2x_1x_3 + \dots 2x_{(n-1)}x_n)}{N} \end{aligned} \quad (9:9)$$

But by definition  $\frac{\sum x_1^2}{N} = \sigma_1^2$ , etc., and  $\frac{\sum x_1x_2}{N} =$  covariance

$x_1x_2 = S_{x_1x_2}$ , etc. The variance of a composite is therefore equal to the sum of the variances of the components plus twice the sum of all possible covariances:

$$\sigma_C^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \dots \sigma_{x_n}^2 + 2S_{x_1x_2} + 2S_{x_1x_3} + \dots 2S_{x_{(n-1)}x_n} \quad (9:10)$$

But from (9:5)  $2S_{x_1x_2} = 2r_{x_1x_2}\sigma_{x_1}\sigma_{x_2}$  (9:10) can be written as:

$$\sigma_C^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \dots \sigma_{x_n}^2 + 2r_{x_1x_2}\sigma_{x_1}\sigma_{x_2} + \dots 2r_{x_{(n-1)}x_n}\sigma_{x_{(n-1)}}\sigma_{x_n}$$

If the variables are weighted the variance can again be worked out fairly simply by using a square table.

	$W_1x_1$	$W_2x_2$	$W_3x_3$	...	$W_nx_n$
$W_1x_1$	$W_1^2x_1^2$	$W_1W_2x_1x_2$	$W_1W_3x_1x_3$	...	$W_1W_nx_1x_n$
$W_2x_2$	$W_1W_2x_1x_2$	$W_2^2x_2^2$	$W_2W_3x_2x_3$	...	$W_2W_nx_2x_n$
$W_3x_3$	$W_1W_3x_1x_3$	$W_2W_3x_2x_3$	$W_3^2x_3^2$	...	$W_3W_nx_3x_n$
...	...	...	...	...	...
$W_nx_n$	$W_1W_nx_1x_n$	$W_2W_nx_2x_n$	$W_3W_nx_3x_n$	...	$W_n^2x_n^2$

The variance of a composite of weighted components will therefore be:

$$\begin{aligned} \sigma_c^2 = & W_1^2\sigma_{x_1}^2 + W_2^2\sigma_{x_2}^2 + \dots W_n^2\sigma_{x_n}^2 + 2W_1W_2S_{x_1x_2} \\ & + \dots 2W_{(n-1)}W_nS_{x_{(n-1)}x_n} \end{aligned} \quad (9:12)$$

For many purposes it is useful to look at the variances of composite variables slightly differently. Such variances are made up of a variance for each of  $n$  tests and a number of terms of the type  $2r_{x_i x_j} \sigma_i \sigma_j$ . The sum of the variances will be  $\sum \sigma^2$ . and the mean variance will be  $\frac{\sum \sigma^2}{N} = \bar{\sigma}^2$ . From this it follows that  $\sum \sigma^2 = n\bar{\sigma}^2$ . Similarly  $\sum r_{ij} \sigma_i \sigma_j = n(n-1)\overline{r_{ij} \sigma_i \sigma_j}$ , as there are  $n(n-1)$  covariance terms. This number is simply the number of terms in the  $n \times n$  table minus the number of variance terms  $= n^2 - n = n(n-1)$ . Hence:

$$\sigma_c^2 = n\bar{\sigma}^2 + n(n-1)\overline{r_{ij} \sigma_i \sigma_j} \quad (9:13)$$

If the component scores were in Z score form the  $\sigma$ 's would disappear and in Z score terms (9:13) would become:

$$\sigma_C^2 = n + n(n-1)\overline{r_{ij}} \quad (9:14)$$

(formula in terms of Z score components).

#### 4. Correlation of a Composite Variable with an Outside Variable

Recalling the formula for  $r_{xy}; r_{xy} = \frac{\sum xy}{N\sigma_x\sigma_y}$ , it can be seen that the correlation between a composite variable © and an outside variable (0) will be:

$$r_{0c} = \frac{\sum x_0x_c}{N\sigma_0\sigma_c} = \frac{\sum x_0(x_1 + x_2 + \dots x_n)}{N\sigma_0\sigma_c} \quad (9:15)$$

This equals:

$$\frac{\sum x_0x_1 + \sum x_0x_2 + \dots \sum x_0x_n}{N\sigma_0\sigma_c}$$

which in turn equals:

$$\frac{\frac{1}{N}(\sum x_0x_1 + \sum x_0x_2 \dots \sum x_0x_n)}{\sigma_0\sigma_c}$$

The terms in the numerator are now all covariance terms and the above can therefore be written (following (9:5)), as:

$$\frac{\sigma_0\sigma_1r_{01} + \sigma_0\sigma_2r_{02} + \sigma_0\sigma_n r_{0n}}{\sigma_0\sigma_c}$$



Dividing by  $\sigma_0$  gives:

$$r_{0c} = \frac{\sigma_1 r_{01} + \sigma_2 r_{02} + \dots + \sigma_n r_{0n}}{\sigma_c} \quad (9:16)$$

In Z score terms this becomes (finding  $\sigma_c$  as the square root of (9:13)):

$$r_{0c} = \frac{\sum r_{oi}}{\sqrt{n + n(n-1)\overline{r_{ij}}}} = \frac{n\overline{r_{oi}}}{\sqrt{n + n(n-1)\overline{r_{ij}}}} \quad (9:17)$$

For reasons which will become apparent in a moment it is convenient to divide numerator and denominator by  $n$  giving:

$$r_{0c} = \frac{\overline{r_{oi}}}{\sqrt{\frac{n}{n^2} + \frac{n(n-1)}{n^2}\overline{r_{ij}}}} = \frac{\overline{r_{oi}}}{\sqrt{\frac{1}{n} + \frac{n-1}{n}\overline{r_{ij}}}} \quad (9:18)$$

As  $n$  increases in size, i.e. as the number of components increases, the value of  $\frac{1}{n}$  becomes smaller and smaller, and  $\frac{n-1}{n}$  becomes nearer and nearer to 1. So with a large number of components, as  $n$  approaches infinity:

$$r_{0c} = \frac{\overline{r_{oi}}}{\sqrt{\overline{r_{ij}}}}; n \rightarrow \infty \quad (9:19)$$

In words the correlation between a composite variable and an outside variable is equal to the mean correlation between components and the outside variable, divided by the square root of the mean intercorrelation between components.

## 5. Correlation between two Composite Variables

Suppose that  $C_X = (X_1 + X_2 + \dots X_n)$  and  $C_Y = (Y_1 + Y_2 + \dots Y_m)$ ,

And that there are  $n$  components in  $C_x$  and  $m$  components in  $C_y$ .

$$r_{C_X C_Y} \text{ will be } \frac{\sum (x_1 + x_2 + \dots x_n)(y_1 + y_2 + \dots y_m)}{N \sigma_{C_X} \sigma_{C_Y}}$$

By steps which will by now be familiar this becomes firstly:

$$\frac{\sum x_1 y_1 + \sum x_2 y_1 + \dots \sum x_n y_m}{N \sigma_{C_X} \sigma_{C_Y}} \quad (9:20)$$

In turn this becomes:

$$\frac{r_{x_1 y_1} \sigma_{x_1} \sigma_{y_1} + r_{x_1 y_2} \sigma_{x_1} \sigma_{y_2} + \dots r_{x_n y_m} \sigma_{x_n} \sigma_{y_m}}{\sigma_{C_X} \sigma_{C_Y}} \quad (9:21)$$

As  $\sum r_{x_i y_i} \sigma_{x_i} \sigma_{y_i} = nm \overline{r_{x_i y_i} \sigma_{x_i} \sigma_{y_i}}$  (9:21) becomes :

$$\frac{nm \overline{r_{x_i y_i} \sigma_{x_i} \sigma_{y_i}}}{\sqrt{n \overline{\sigma_{x_i}^2} + n(n-1) \overline{r_{x_i x_j} \sigma_{x_i} \sigma_{x_j}}} \sqrt{m \overline{\sigma_{y_i}^2} + m(m-1) \overline{r_{y_i y_j} \sigma_{y_i} \sigma_{y_j}}}} \quad (9:22)$$

If all components are in  $Z$  score form, this becomes:

$$\frac{nm \overline{r_{x_i y_i}}}{\sqrt{n + n(n-1) \overline{r_{x_i x_j}}} \sqrt{m + m(m-1) \overline{r_{y_i y_j}}}} \quad (9:23)$$

Dividing numerator and denominator by  $nm$  gives:

$$\frac{\overline{r_{x_i y_i}}}{\sqrt{\frac{n}{n^2} + \frac{n(n-1)}{n^2} \overline{r_{x_i x_j}}}} \sqrt{\frac{m}{m^2} + \frac{m(m-1)}{m^2} \overline{r_{y_i y_j}}} \quad (9:24)$$

$$= \frac{\overline{r_{x_i y_i}}}{\sqrt{\frac{1}{n} + \frac{n-1}{n} \overline{r_{x_i x_j}}}} \sqrt{\frac{1}{m} + \frac{m-1}{m} \overline{r_{y_i y_j}}}$$

As the number of components in each composite becomes larger  $\frac{1}{n}$  and  $\frac{1}{m}$  become closer to zero, and  $\frac{n-1}{n}$  and  $\frac{m-1}{m}$  become nearer to

1. Therefore, as  $n$  and  $m$  approach infinity we obtain:

$$\frac{\overline{r_{x_i y_i}}}{\sqrt{\overline{r_{x_i x_j}} \overline{r_{y_i y_j}}}} \quad (9:25)$$

### Problems

- A. Suppose a composite is formed of three tests  $X_1$ ,  $X_2$ , and  $X_3$ , each with a mean of 10 and a standard deviation of 3. If  $r_{x_1 x_2} = .30$ ,  $r_{x_1 x_3} = .40$ , and  $r_{x_2 x_3} = .50$ , what will be the mean and variance of the composite scores?
- B. A composite is made up of four tests  $X_1, X_2, X_3$ , and  $X_4$ , with  $M_{x_1} = 10$ ,  $M_{x_2} = 20$ ,  $M_{x_3} = 30$ ,  $M_{x_4} = 40$ ; and  $\sigma_{x_1} = 2$ ,  $\sigma_{x_2} = 3$ ,  $\sigma_{x_3} = 4$ , and  $\sigma_{x_4} = 5$ . If  $r_{x_1 x_2} = .20$ ,  $r_{x_1 x_3} = .30$ ,  $r_{x_1 x_4} = .30$ ,  $r_{x_2 x_3} = .40$ ,  $r_{x_2 x_4} = .40$ , and  $r_{x_3 x_4} = .20$ , what is the mean of the composite and what is its variance?
- C. Given a weighted composite  $C = 2X_1 + 3X_2$ , and  $r_{12} = .40$ ;  $M_{x_1} = 10$ , and  $M_{x_2} = 20$ ; and  $\sigma_{x_1} = 5$  and  $\sigma_{x_2} = 6$ , what is the mean of  $C$  and what is its variance?

- D. Given two composites X and Y what will be their correlation given the following data on the composites:

	$X_2$	$Y_1$	$Y_2$	$Y_3$	$M$	$\sigma$
$X_1$	.20	.10	.20	.30	10	3
$X_2$		.20	.30	.40	15	4
$Y_1$			.50	.60	20	5
$Y_2$				.40	25	5
$Y_3$					30	5

Use Formula (9:21).

### Answers

A.  $\bar{C} = \bar{X}_1 + \bar{X}_2 + \bar{X}_3 = 10 + 10 + 10 = 30$

$$\begin{aligned}\sigma_C^2 &= \sigma_{X_1}^2 + \sigma_{X_2}^2 + \sigma_{X_3}^2 + 2r_{X_1X_2} \sigma_{X_1} \sigma_{X_2} + 2r_{X_1X_3} \sigma_{X_1} \sigma_{X_3} + 2r_{X_2X_3} \sigma_{X_2} \sigma_{X_3} \\ &= 9 + 9 + 9 + 2(.30 \times 3 \times 3) + 2(.40 \times 3 \times 3) + 2(.50 \times 3 \times 3) \\ &= + 48.6.\end{aligned}$$

B.  $\bar{C} = 10 + 20 + 30 + 40 = 100$

$$\begin{aligned}\sigma_C^2 &= 4 + 9 + 16 + 25 + 2(.20 \times 2 \times 3) + 2(.30 \times 2 \times 4) \\ &\quad + 2(.30 \times 2 \times 5) + 2(.40 \times 3 \times 4) + 2(.40 \times 3 \times 5) + 2(.20 \times 4 \times 5) \\ &= 96.8.\end{aligned}$$

$$C. \quad \bar{C} = W_1\bar{X}_1 + W_2\bar{X}_2 = (2 \times 10) + (3 \times 20) = 80$$

$$\begin{aligned}\sigma_C^2 &= W_1^2\sigma_{X_1}^2 + W_2^2\sigma_{X_2}^2 + 2W_1W_2r_{X_1X_2}\sigma_{X_1}\sigma_{X_2} \\ &= (4 \times 25) + (9 \times 36) + (2 \times 2 \times 3 \times .40 \times 5 \times 6) \\ &= 568.\end{aligned}$$

D.

$$r_{C_xC_y} = \frac{r_{X_1Y_1}\sigma_{X_1}\sigma_{Y_1} + r_{X_1Y_2}\sigma_{X_1}\sigma_{Y_2} + r_{X_1Y_3}\sigma_{X_1}\sigma_{Y_3} + r_{X_2Y_1}\sigma_{X_2}\sigma_{Y_1} + r_{X_2Y_2}\sigma_{X_2}\sigma_{Y_2} + r_{X_2Y_3}\sigma_{X_2}\sigma_{Y_3}}{\sigma_{C_x}\sigma_{C_y}}$$

$$\begin{aligned}\text{The numerator} &= (.10 \times 3 \times 5) + (.20 \times 3 \times 5) \\ &\quad + (.30 \times 3 \times 5) + (.20 \times 4 \times 5) \\ &\quad + (.30 \times 4 \times 5) + (.40 \times 4 \times 5) \\ &= 27.0\end{aligned}$$

The denominator =

$$\begin{aligned}\sqrt{\sigma_{C_x}^2} &= \sqrt{9 + 16 + 2(.20 \times 3 \times 4)} = \sqrt{29.8} \\ \sqrt{\sigma_{C_y}^2} &= \sqrt{\frac{25 + 25 + 25 + 2(.50 \times 5 \times 5) + 2(.60 \times 5 \times 5)}{+ 2(.40 \times 5 \times 5)}} \\ &= \sqrt{150}\end{aligned}$$

$$\text{Therefore } r_{C_xC_y} = \frac{27.0}{\sqrt{29.8}\sqrt{150}} = .40$$

## 6. The Multiple Regression Equation Again

On reflection it will by now be clear that in multiple regression, the regression weights combined appropriately with the raw scores or Z scores yield weighted composite scores. These scores when correlated with the criterion scores result in the multiple correlation coefficient. Thus  $R_{0.12...n}$ , although it expresses the relationship between a number of variables and the criterion, is in fact the correlation between only two variables - the weighted composite variable and the criterion.

Thus the mean score on the predicted variable and its variance can be worked out from the formulae for weighted composites. It is merely necessary to substitute  $\beta$ 's or  $b$ 's for the  $w$ 's in Formulae 9:2 and 9:12. It can be shown that the mean of the composite will equal the mean of the criterion.

$$\hat{M}_0 = M_0 \quad (9:26)$$

Where:

$\hat{M}_0$  = the mean of the predicted scores, and  
 $M_0$  = the mean of the criterion scores

*Proof*

$$(1) \quad \hat{X}_0 = M_0 - b_1M_1 - \dots - b_nM_n + b_1X_1 + \dots + b_nX_n$$

$$(2) \quad \hat{M}_0 = \frac{\sum (M_0 - b_1M_1 - \dots - b_nM_n + b_1X_1 + \dots + b_nX_n)}{N}$$

$$(3) \quad = \frac{NM_0 - b_1NM_1 - \dots - b_nNM_n + b_1\sum X_1 + \dots + b_n\sum X_n}{N}$$

$$(4) \quad = M_0 - b_1M_1 - \dots - b_nM_n + b_1M_1 + \dots + b_nM_n$$

$$(5) \quad = M_0$$

The variance of the predicted scores will be:

$$\hat{\sigma}_0^2 = \sum b_i^2 \sigma_i^2 + 2 \sum b_i b_j \sigma_i \sigma_j r_{ij} \quad (9:27)$$

*Proof*

Once more a table will help in working out the products of the weighted deviation scores:

	$b_1x_1$	$b_2x_2$	...	$b_nx_n$
$b_1x_1$	$b_1^2x_1^2$	$b_1x_1b_2x_2$	...	$b_1x_1b_nx_n$
$b_2x_2$	$b_1x_1b_2x_2$	$b_2^2x_2^2$	...	$b_2x_2b_nx_n$
...	...	...	...	...
$b_nx_n$	$b_1x_1b_nx_n$	$b_2x_2b_nx_n$	...	$b_n^2x_n^2$

Summing these values across individuals will give:

$$(1) \quad \sum \hat{x}_0^2 = b_1^2 \sum x_1^2 + b_2^2 \sum x_2^2 + \dots + b_n^2 \sum x_n^2 + 2b_1b_2 \sum x_1x_2 + \dots + 2b_{(n-1)}b_n \sum x_{(n-1)}x_n$$

(2) Dividing by  $N$  gives:

$$b_1^2 \sigma_1^2 + b_2^2 \sigma_2^2 + \dots + b_n^2 \sigma_n^2 + 2b_1b_2 \sigma_1 \sigma_2 r_{12} + \dots + 2b_{(n-1)}b_n \sigma_{(n-1)} \sigma_n$$

(3) This equals:

$$\sum b_i^2 \sigma_i^2 + 2 \sum b_i b_j \sigma_i \sigma_j r_{ij}$$

*Problems*

- A. If  $r_{01} = .30$ ,  $r_{02} = .40$ , and  $r_{12} = .50$ , write the Z score regression equation for predicting 0 from 1 and 2.
- B. What will be the mean of the predicted scores? (Use the formula for the mean of a weighted composite.)
- C. What will the variance of the predicted scores be? (Use the formula for the variance of a weighted composite.)

*Answers*

A.  $Z_0 = \beta_1 Z_1 + \beta_2 Z_2; \beta_1 = \frac{r_{01} - r_{02} r_{12}}{1 - r_{12}^2}$

and  $\beta_2 = \frac{r_{02} - r_{01} r_{12}}{1 - r_{12}^2}$

Therefore  $\beta_1 = \frac{.30 - (.40 \times .50)}{1 - .50^2} = .13$

$$\beta_2 = \frac{.40 - (.30 \times .50)}{1 - .50^2} = .33$$

Therefore  $Z_0 = .13Z_1 + .33Z_2$

B.  $\hat{M}_{Z_0} = .13M_{Z_1} + .33M_{Z_2} = 0$



C. (1) A square table will help:

	$\beta_1 Z_1$	$\beta_2 Z_2$
$\beta_1 Z_1$	$\beta_1^2 Z_1^2$	$\beta_1 \beta_2 Z_1 Z_2$
$\beta_2 Z_2$	$\beta_2 \beta_1 Z_2 Z_1$	$\beta_2^2 Z_2^2$

(2) Summing across individuals and dividing by  $N$  gives:

$$\hat{\sigma}_{z_0}^2 = \beta_1^2 \frac{\sum Z_1^2}{N} + \beta_2^2 \frac{\sum Z_2^2}{N} + 2\beta_1 \beta_2 \frac{\sum Z_1 Z_2}{N}$$

(3) As  $\frac{\sum Z_1^2}{N}$  and  $\frac{\sum Z_2^2}{N}$  equal 1; and as  $\frac{\sum Z_1 Z_2}{N}$  equals  $r_{12}$  we obtain:

$$\begin{aligned} (4) \quad \hat{\sigma}_{z_0}^2 &= \beta_1^2 + \beta_2^2 + 2\beta_1 \beta_2 r_{12} \\ &= .13^2 + .33^2 + (2 \times .13 \times .33 \times .50) \\ &= .17 \end{aligned}$$

(5) Therefore  $\hat{\sigma}_{z_0} = .41$ .

According to (8:13)  $\hat{\sigma}_0 = R_{0.12} \sigma_0$ .

Does this agree with the answer just obtained?