Questionnaire Construction: Items

**Item Content**
- Judgments - refer to items where there is a correct response.
- Sentiments - refer to items that elicit responses about personal reactions, preferences, interests, attitudes, values, and likes and dislikes.

**Item Response Categories**
- Closed Ended (Fixed Alternative) - These items allow the subject to select from one or more categories provided by the questionnaire.
- Open Ended - These items do not specify response categories.

**Closed Ended Items - Advantages**
- Comparability of answers.
- Easier to code.
- Item meaning clearer.
- Answers tend to be complete and relevant.
- Greater inclination to answer sensitive questions.
- Easier to answer.

**Closed Ended Items - Disadvantages**
- Guessing or random answering.
- Inappropriate/irrelevant response categories.
- Too many categories.
- Undetected differences in interpretation.
- Less variation in answers.
- Higher likelihood of clerical error.
**Open Ended Items - Advantages**

- Can use when all possible response categories are not known.
- Allow for more detail, clarification, and qualification.
- Can be used when there are too many potential answer categories to list.
- Preferable for complex issues.
- Allow more opportunity for exploration & self expression.

**Open Ended Items - Disadvantages**

- May lead to collection of worthless & irrelevant information.
- Data are not standard from person to person.
- Low intercoder reliability.
- Require superior respondent writing skills.
- Questions may be too general for respondent to understand.

**Item Construction Guidelines**

- Avoid double barreled questions.
- Make items clear.
- Avoid biased items and terms.
- Avoid long questions.
- Make sure that respondents are competent to answer questions.

**Item Construction Guidelines - 2**

- Break up response set opportunities.
- Avoid negative items.
- Provide an explicit middle.
- Carefully word threatening questions

**Correlation and Regression**

**Slope and Intercept**

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
</tr>
</tbody>
</table>
Slope and Intercept

- Slope = 2.0
- Y-intercept = 7

The Regression (Prediction) Line

- \( Y_{\text{pred}} = a + bX \)
- \( Y_{\text{pred}} \) is the predicted value of \( Y \)
- \( a \) is the Y-intercept, and
- \( b \) is the slope
- \( E = Y - Y_{\text{pred}} \)
- \( Y = Y_{\text{pred}} + E \)
- \( Y = a + bX + E \)
**Error and Prediction**

\[ E = Y - Y_{\text{pred}} \]

- When \( X = 2 \), \( Y = 4 \) and \( Y = 5 \).
  - The predicted value of \( Y \) is
    \[ Y_{\text{pred}} = 3.5 + 0.5(2) = 3.5 + 1.0 = 4.5 \]
  - The prediction errors for the two values of \( Y \) are
    - \( Y = 4 \): \( E = Y - Y_{\text{pred}} = 4 - 4.5 = -0.5 \)
    - \( Y = 5 \): \( E = Y - Y_{\text{pred}} = 5 - 4.5 = +0.5 \)

**Error Sum of Squares**

\[ SSE = \Sigma (E)^2 = \Sigma (Y - Y_{\text{pred}})^2 \]

**Slope and Intercept - Computational Formulae**

- **Slope**
  \[ b = \frac{SS_{XY}}{SS_X} \]
  - where
    \[ SS_{XY} = \Sigma (X - \bar{X})(Y - \bar{Y}), \quad SS_X = \Sigma (X - \bar{X})^2 \]
  - **Y Intercept**
    \[ a = \bar{Y} - b\bar{X} \]

**Computing the Regression Coefficients**

<table>
<thead>
<tr>
<th>( Y )</th>
<th>( X )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

**Computing the Regression Coefficients**

<table>
<thead>
<tr>
<th>( Y )</th>
<th>( Y - \bar{Y} )</th>
<th>( (Y - \bar{Y})^2 )</th>
<th>( X )</th>
<th>( X - \bar{X} )</th>
<th>( (X - \bar{X})^2 )</th>
<th>( (Y - \bar{Y})(X - \bar{X}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-4</td>
<td>16</td>
<td>1</td>
<td>-2</td>
<td>4</td>
<td>(+4)(-2) = -8</td>
</tr>
<tr>
<td>2</td>
<td>-2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>(-2)(0) = 0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>(0)(-1) = 0</td>
</tr>
<tr>
<td>6</td>
<td>+2</td>
<td>4</td>
<td>5</td>
<td>+2</td>
<td>4</td>
<td>(+2)(+2) = +4</td>
</tr>
<tr>
<td>8</td>
<td>+4</td>
<td>16</td>
<td>4</td>
<td>+1</td>
<td>1</td>
<td>(+4)(+1) = +4</td>
</tr>
</tbody>
</table>

\[ SS_y = 40 \quad SS_x = 10 \quad SS_{xy} = +16 \]

- \( SS_X = \Sigma (X - \bar{X})^2 = \Sigma (X - \bar{X})(X - \bar{X}) \)
- \( SS_{XY} = \Sigma (X - \bar{X})(Y - \bar{Y}) \)
Computing the Regression Coefficients

**Slope**

\[ b = \frac{SS_{XY}}{SS_X} = \frac{+16}{10} = +1.6 \]

**Y Intercept**

\[ a = \overline{Y} - b \overline{X} = 4 - (+1.6)(3) = 4 - 4.8 = -0.8 \]

---

Computing the Regression Coefficients: Exercise

<table>
<thead>
<tr>
<th>Y</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

---

Computing the Regression Coefficients: Setup

<table>
<thead>
<tr>
<th>Y</th>
<th>Y - ( \overline{Y} )</th>
<th>( (Y - \overline{Y})^2 )</th>
<th>X</th>
<th>X - ( \overline{X} )</th>
<th>( (X - \overline{X})^2 )</th>
<th>( (Y - \overline{Y})(X - \overline{X}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>16</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>36</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>64</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>36</td>
</tr>
</tbody>
</table>

\[ SS_Y = \quad SS_X = \quad SS_{XY} = \]

---

Sums of Squares

<table>
<thead>
<tr>
<th>Type</th>
<th>Formula</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>( SS_R = \sum(Y_{pred} - \overline{Y})^2 )</td>
<td>A measure of the total variability of predicted score values around the mean</td>
</tr>
<tr>
<td>Error</td>
<td>( SS_E = \sum(Y - Y_{pred})^2 )</td>
<td>A measure of the total variability of obtained score values around their predicted values</td>
</tr>
<tr>
<td>Total</td>
<td>( SS_T = \sum(Y - \overline{Y})^2 )</td>
<td>A measure of the total variability of obtained score values around the mean</td>
</tr>
</tbody>
</table>

\[ Y - \overline{Y} = (Y - Y_{pred}) + (Y_{pred} - \overline{Y}) \]

\[ SS_T = SS_E + SS_R \]
Standard Error of Estimate

- Variance Error
  \[ \text{SSE} \]
  \[ \text{se}^2 = \frac{\text{SSE}}{n - 2} \]

- Standard Error of Estimate
  \[ \text{SE} = \sqrt{(\text{se}^2)} \]

Standard Error of Estimate: Exercise

- Total Sum of Squares = 3884.550
- Regression Sum of Squares = 1413.833
- Error Sum of Squares = 2470.717

- Prediction Equation:
  \[ Y_{\text{pred}} = 9.232 + 0.817X \]

Proportion of Variance Explained (PVE)

\[ \text{PVE} = \frac{\text{SSR}}{\text{SST}} \]
\[ \text{PVE} = \frac{(\text{SS}_{xy})^2}{\text{SS}_x \text{SS}_y} \]

Coefficient of Determination

\[ r^2 = \frac{\text{SSR}}{\text{SST}} \]
\[ 1 - r^2 = 1 - \frac{\text{SSR}}{\text{SST}} = \frac{\text{SSE}}{\text{SST}} \]
**PVE: Exercise**

- Total Sum of Squares = 3884.550
- Regression Sum of Squares = 1413.833
- Error Sum of Squares = 2470.717
- Prediction Equation:
  \[ Y_{\text{pred}} = 9.232 + 0.817X \]
- Standard Error of Estimate = 11.72

---

**Correlation and Regression**

The Correlation Coefficient

\[
\begin{align*}
    r &= \frac{SS_{XY}}{\sqrt{SS_X SS_Y}} \\
    b &= \frac{SS_{XY}}{SS_X} \\
    \end{align*}
\]

Computing the Correlation Coefficients

\[
\begin{array}{cccccccc}
    Y & Y - \bar{Y} & (Y - \bar{Y})^2 & X & X - \bar{X} & (X - \bar{X})^2 & (Y - \bar{Y})(X - \bar{X}) \\
    \hline
    0 & -4 & 16 & 1 & -2 & 4 & (-4)(-2) = +8 \\
    2 & -2 & 4 & 3 & 0 & 0 & (-2)(0) = 0 \\
    4 & 0 & 0 & 2 & -1 & 1 & (0)(-1) = 0 \\
    6 & +2 & 4 & 5 & +2 & 4 & (+2)(+2) = +4 \\
    8 & +4 & 16 & 4 & +1 & 1 & (+4)(+1) = +4 \\
    \hline
    \text{SS}_Y &= 40 & \text{SS}_X &= 10 & \text{SS}_{XY} &= +16 \\
    \hline
\end{array}
\]

\[
\begin{align*}
    r &= \frac{SS_{XY}}{\sqrt{SS_X SS_Y}} = \frac{+16}{\sqrt{(10)(40)}} = \frac{+16}{\sqrt{400}} = \frac{+16}{20} = +0.8 \\
\end{align*}
\]
Correlation

\[ Y_{pred} = -0.8 + 1.6X \]

Standard Error of Estimate
(The Easy Way)

\[ S_E = s_Y \sqrt{(1 - r^2) \frac{n-1}{n-2}} \]

Spearman Rho (\( r_s \))

- Used when X and/or Y are ordinal variables
- Procedure
  - Assign ranks to X \( R_X \)
  - Assign ranks to Y \( R_Y \)
  - For each pair, compute \( d = R_Y - R_X \)
  - \( r_s = 1 - \frac{6 \Sigma d}{n(n^2-1)} \)

Phi (\( \phi \)) Coefficient

- Used when X and Y are nominal variables
- Procedure
  \[ \phi = \frac{(BC) - (AD)}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} \]

Applied Measurement Theory

Obtained, True, and Error Scores

\[ X = T + E \]

- \( X \) is the observed score
- \( T \) is the true score, and
- \( E \) is the error score.
Evaluating Reliability -
Score Variance

\[ \sigma_X^2 = \sigma_T^2 + \sigma_E^2 \]

The Reliability Coefficient

\[ r_{kk} = r_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} \]

Procedures for Evaluating
Reliability

- Retest (Stability)
- Parallel Forms (Equivalence)
- Internal Consistency (Item Homogeneity)

Retest Reliability

- \( r_{kk} = r_{1st,2nd} \)
- One group of people
- One testing procedure (instrument)
- Two measurement times.

Parallel Forms Reliability

- \( r_{kk} = r_{form a, form b} \)
- One group of people
- Two testing procedures (instruments)
- One measurement time.

Internal Consistency Reliability

- Split Half Estimation
- Coefficient Alpha (Cronbach’s Alpha)
- One group of people
- One testing procedures (instruments)
- One measurement time.