cuantil de una distribución $N(0,1)$ correspondiente a probabilidad $1-\alpha/2$:

$$\int_{-\infty}^{Z_{1-\alpha/2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx = 1-\alpha/2 .$$

$$1-\alpha \overset{\text{aprox}}{\approx} \mathbb{P}\left(\left|\frac{\hat{\psi}_q-\psi_q}{w/\sqrt{n-1}}\right| \le Z_{1-\alpha/2}\right)$$

$$= \mathbb{P}\left(-Z_{1-\alpha/2} \le \frac{\hat{\psi}_q-\psi_q}{w/\sqrt{n-1}} \le Z_{1-\alpha/2}\right)$$

$$= \mathbb{P}\left(-Z_{1-\alpha/2} \le \frac{\psi_q-\hat{\psi}_q}{w/\sqrt{n-1}} \le Z_{1-\alpha/2}\right)$$

$$= \mathbb{P}\left(\hat{\psi}_q - \frac{w}{\sqrt{n-1}}\times Z_{1-\alpha/2} \le \psi_q \le \hat{\psi}_q + \frac{w}{\sqrt{n-1}} * Z_{1-\alpha/2}\right).$$

Entonces, a nivel $(1-\alpha)\times 100\%$ de confianza ψ_q debe estar en el intervalo

$$\left(\hat{\psi}_q - \frac{w}{\sqrt{n-1}}\times Z_{1-\alpha/2},\ \hat{\psi}_q + \frac{w}{\sqrt{n-1}}\times Z_{1-\alpha/2}\right).$$

Para los datos de billetes del banco suizo, sabemos que $\hat{\psi}_1 = \frac{l_1}{l_1+l_2+\cdots+l_6} = 0.67$.

Supóngase que quisieramos probar la hipótesis (h).... $\begin{cases} H_0: \psi_1 = 0.75 \\ \quad \text{vs.} \\ H_1: \text{No } H_0 \end{cases}$

Una forma de llevar a cabo la prueba de hipótesis es usando intervalos de confianza

$$\hat{\beta} = \frac{l_1^2}{l_1^2 + \cdots + l_6^2} = \frac{(2.985)^2}{(2.985)^2 + (0.931)^2 + \cdots + (0.035)^2}$$

$$= 0.902 .$$

$$\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_6^2 = tr(\Lambda^2) = tr(\Lambda^2 \Gamma' \Gamma)$$
$$= tr(\Gamma \Lambda^2 \Gamma')$$
$$= tr(\Gamma \Lambda \Gamma' \Gamma \Lambda \Gamma') = tr(\Sigma \Sigma)$$
$$= tr(\Sigma^2)$$

Se puede estimar con $tr(\hat{\Sigma}^2) = \sum_{j=1}^{6} l_j^2$

$$= 9.883$$

$tr(\Sigma)$ se puede estimar con $tr(\hat{\Sigma})$

$$= tr(\mathcal{G} \mathcal{L} \mathcal{G}') = tr(\mathcal{L} \mathcal{G}' \mathcal{G}) = tr(\mathcal{L})$$

$$= l_1 + \cdots + l_6 = 4.472$$

para $q=1$

De manera que w en la ecuación (u), página 63 se puede estimar por

$$\hat{w}^2 = \frac{2\,tr(\hat{\Sigma}^2)}{[tr(\hat{\Sigma})]^2}\left(\hat{\Psi}_1^2 - 2\hat{\psi}_1\hat{\beta} + \hat{\beta}\right)$$

$$= \frac{2(9.883)}{(4.472)^2}\left[(0.668)^2 - 2(0.668)(0.902) + 0.902\right]$$

$$= 0.142 .$$

Entonces, a nivel $0.95 \times 100\%$ ($\alpha = 0.05$) de significancia un intervalo (aproximado) para ψ_1 es (véase página 64)

$$\left(0.668 - \sqrt{\frac{0.142}{199}} \times 1.96 ,\ 0.668 + \sqrt{\frac{0.142}{199}} \times 1.96\right)$$

$$= (0.615, 0.720),$$

el cual no contiene el valor de 0.75, por tanto rechazamos Ho en (h).

ANALISIS DE COMPONENTES PRINCIPALES NORMALIZADO

Como se mencionó antes, las componentes principales son sensibles a cambios de escala en las variables, de esta forma al re-escalar alguna de las variables bajo estudio los resultados pueden cambiar.

Por otra parte, uno de los objetivos en el análisis de componentes principales es encontrar las direcciones en $\mathbb{R}^p$, para las cuales los datos presentan mayor variabilidad. Sin embargo si los datos originales presentan heterogeneidad con respecto a sus varianzas, por ejemplo si la escala de la variable X_i es en Kilogramos y la escala de X_j este en dólares, entonces "la dirección en $\mathbb{R}^p$ para la cual los datos presentan mayor variabilidad" podría originarse por la heterogeneidad en las escalas y no debido a la naturaleza de los datos.

Para evitar estos problemas se utiliza una estandarización de las variables en el vector $\mathcal{X}$, a saber