STATISTICS FOR ENGINEERS
STAT4073 :: PROBLEM 3.X

Extension: Frank D. Lucas (R-OK3) is the US Representative for Oklahoma's Third Congressional district, which includes all of Stillwater. Lucas has held this seat for 18 years. He faced no primary challenge from his fellow Republicans and will face Democrat Frankie Robbins in a rematch in November. Two years ago, Lucas beat Robbins 70% to 24%.

Recent polling (scant though it is) suggests that the results may be different. Rasmussen polled likely voters and concluded the support for Lucas in his district is $80\% \pm 5\%$. Gallup did likewise and found support at $85\% \pm 3\%$. A start-up polling firm, CEF, concluded that support for Lucas was at $90\% \pm 8\%$.

Your client has hired you to determine the actual support for Lucas (along with the errors). Do the calculations separately, not to be handed in. In a memo to your client, Sam Warde, summarize your technique (in a manner that is both readable and correct) and provide the best estimate of the support for Lucas in OK-3. Include both the estimate and the error range.

Solution: This problem is an extension to an example we did in class. In that example, there were only two polls available, so the weighting was c and $1 - c$. Afterwards, I mentioned that to handle more polls, one would have to weight according to c, d, and $1 - c - d$, as we will here.

The overview of the steps is to create the formula for the variance, minimize it (thus getting the weights), and calculate the weighted averages.

Let us assign c to the weight to be given to the Rasmussen poll; d to the Gallup poll; and $(1-c-d)$ to the CEF poll, then the variance of the final estimate will be

$$\sigma^2 = 25c^2 + 9d^2 + 64(1-c-d)^2 \tag{1}$$

To minimize the variance, one needs to differentiate Eqn 1 first with respect to c, then with respect to d. You will then have two equations with two unknowns, which is solvable.

$$\begin{aligned} \frac{\partial \sigma^2}{\partial c} =& 50c + 0 + -128(1-c-d) \\ \Rightarrow 0 =& 50c + 0 + -128(1-c-d) \\ 178c =& 128 - 128d \\ c =& \frac{128-128d}{178} \end{aligned} \tag{2}$$

Similarly for d, we get

$$d = \frac{128-128c}{146}$$

Thus, solving first for d, then for c, we have:

$$\begin{aligned} d =& \frac{128 - 128\left(\frac{128-128d}{178}\right)}{146} \\ 146d =& 128 - \frac{128^2}{178}\Big(128 - 128d\Big) \\ \left(146 - \frac{128^2}{178}\right) d =& 128 - \frac{128^2}{178} \end{aligned}$$

This gives us $d \approx 0.666389$. Thus, from Eqn 2, we get $c \approx 0.2399$.

Now, we take those weights and use them to determine the best estimate of the support for Lucas:

$$X_p = 80c + 85d + 90(1-c-d) \approx 84.26905,$$

and of the error:

$$\sigma_p^2 = 25c^2 + 9d^2 + 64(1 - c - d)^2 \approx 5.997501$$

Thus, the best estimate of the support for Lucas is $84\% \pm 2\%$.

◇