

## Math 2a Solution Set 5

**Problem 1**

Once a sample is chosen, its mean is a fixed number. But the sample itself is random! That is to say that the randomness of the sample mean is given by the randomness of the choice of the sample. This randomness manifests itself in the different values that the mean (and, for that matter, any other statistic that we may want to consider) takes for different (random) choices of samples from our population.

**Problem 2**

We know (Moore and McCabe, page 376) that when we draw a simple random sample of size  $n$  from a large population with population proportion of successes  $p$  (in this case families in a certain area who are living below the poverty level) and we consider  $\hat{p}$  the sample proportion of successes, then  $\hat{p}$  is distributed like a normal distribution with mean  $p$  and std  $\sqrt{p(1-p)/n}$ . Moreover the problem tells us that the population proportion is  $p = .15$ . We want to find  $n$  so that standard error of the estimate is  $\sqrt{p(1-p)/n} = .02$ . Since  $\frac{p(1-p)}{(.02)^2} = \frac{.15 \cdot .85}{(.02)^2} = 318.75$  we get  $n = 319$ .

You may also interpret the problem in this way (following your text): We know that the standard error  $SE_{\hat{p}} = \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} = .02$ . Moreover the problem tells us that the proportion is  $.15$ , and so we may assume  $\tilde{p} = .15$ . Since  $\frac{\tilde{p}(1-\tilde{p})}{(.02)^2} = \frac{.15 \cdot .85}{(.02)^2} = 318.75$ , we get  $n + 4 = 319 \Rightarrow n = 315$ . As announced in class, we accept both answers.

**Problem 5.66**

- (a) The sample mean  $\bar{x}$  of an SRS of size  $n = 23$  from a large population (the population of all students using the control method of studying Russian) distributed according to  $N(32, 6)$  is itself distributed according to  $N(32, \frac{6}{\sqrt{23}})$ . Similarly, since the distribution of the population of students whose oral practice is delayed is  $N(29, 5)$ , the sampling distribution of the mean score  $\bar{y}$  in the experimental group is  $N(29, \frac{5}{\sqrt{23}})$ .
- (b) As the two SRS are independent, we know that  $\bar{x}$  and  $\bar{y}$  are two independent normal distributions. Then their difference  $\bar{y} - \bar{x}$  is also a normal distribution with mean and variance

$$\mu_{\bar{y}-\bar{x}} = \mu_{\bar{y}} - \mu_{\bar{x}} = 29 - 32 = -3$$

$$\sigma_{\bar{y}-\bar{x}}^2 = \sigma_{\bar{y}}^2 + \sigma_{\bar{x}}^2 = \frac{5^2}{23} + \frac{6^2}{23} = \frac{61}{23} \approx 2.6522$$

- (c) As  $\bar{y} - \bar{x}$  is distributed as  $N(-3, \sqrt{\frac{61}{23}})$ , we have that

$$P(\bar{y} \geq \bar{x}) = P(\bar{y} - \bar{x} \geq 0) = 1 - P(\bar{y} - \bar{x} < 0) \approx 1 - \Phi\left(\frac{0 - (-3)}{\sqrt{\frac{61}{23}}}\right) \approx 1 - \Phi(1.8421) \approx 1 - .9671 = .0329$$

**Problem 6.22**

(a) We have chosen an SRS of size  $n = 5$  and sample mean  $\bar{x} = 10.0023$  from a population having unknown mean  $\mu$  and known standard deviation  $\sigma = .0002$ . Then a 98% confidence interval for  $\mu$  is

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 10.0023 \pm 2.326 \frac{0.0002}{\sqrt{5}} = 10.0023 \pm 0.000208$$

(b) The 98% confidence interval will have a margin of error  $m = 0.0001$  for a sample of size

$$\left( \frac{z^* \sigma}{m} \right)^2 \approx \left( \frac{2.326 \cdot 0.0002}{0.0001} \right)^2 = 21.64 \Rightarrow n = 22$$

**Problem 6.26**

(a) As we assume that we compute confidence intervals based on independent samples from the three states, we get that the median household incomes for the three states are also independent. Thus the probability that all three 95% confidence intervals cover the true median incomes will be the product of the three confidence levels, i.e.  $(.95)^3 \approx .8574 = 85.74\%$ .

(b) Similarly, the probability that at least two of the three intervals cover the true median incomes is the probability that exactly two intervals cover the true medians (and one doesn't) plus the probability that all three 95% confidence intervals cover the true median incomes, i.e.  $3 \cdot (.95)^2 \cdot .05 + (.95)^3 \approx .9928 = 99.28\%$ .