**(1)** [2pts] In <u>2012</u>, what percent had salaries below $120,000? **(D)**

**(2)** [3pts] In <u>2013</u>, what percent had salaries within one standard deviation of the mean? **(E)**

**(3)** [2pts] In <u>2012</u>, what is the interquartile range? **(A)**

**(4)** [3pts] From 2012 to 2013, what percent had their salaries go down (in absolute dollar value)? **(C)**

**(5)** [5pts] What is the coefficient of correlation between salaries in 2012 and 2013? **(E)**

**(6)** [4pts] In the highly unrealistic scenario where there is no relationship between salaries in 2012 and 2013, what would be the standard deviation of the change in salaries from 2012 to 2013? **(D)**

**(7)** [3pts] In the highly unrealistic scenario where each employee's salary in 2013 is simply $2,000 higher than in 2012, what would be the coefficient of correlation between salaries in 2012 and 2013? **(E)**

**(8)** [4pts] If a passing score is 50%, what percent of employees passed the performance review? **(D)**

**(9)** [4pts] To earn a bonus, how high of a score does an employee need? **(A)**

**(10)** [4pts] Roughly, how many households fall in the bin with the tallest bar? **(B)**

**(11)** [4pts] Roughly, what is the standard deviation? **(C)**

**(12)** [4pts] How to interpret the slope coefficient in the <u>Southeast Regression</u>? Within the Southeast, geographic areas where the portion that are black is ___ higher have absolute mobility that is ___ percentiles lower on average. **(E)**

**(13)** [4pts] How to interpret the slope coefficient in the <u>Rest of U.S. Regression</u>? Outside of the Southeast, geographic areas where the portion that are black is ___ higher have absolute mobility that is ___ percent lower on average. **(B)**

**(14)** [4pts] For the <u>Southeast</u>, if both the absolute mobility and fraction black variables are standardized, what would be the slope coefficient for a regression on the *standardized* data? **(B)**

**(15)** [3pts] What is $P(F \mid C')$, which also may be written as $P(F \mid C^c)$? **(C)**

**(16)** [4pts] What is the chance that a flagged teacher is a cheater? **(B)**

**(17)** [4pts] If $X$ is a Binomial random variable, which of these statements are TRUE? **(A)**

**(18)** [5pts] In the incomplete graph below of the sampling distribution of $\hat{P}$ for $n = 155$ and $p = 0.5$, what value goes in the spot marked "**?**" on the x-axis? **(A)**

**(19)** [2pts] If an analyst receives 100 files, what is the expected total hours? **(E)**

**(20)** [5pts] If an analyst receives 100 files, what is the standard deviation of total hours? **(C)**

**(21)** [3pts] For the first five files, what is the chance that *none are trivial*? **(E)**

**(22)** [3pts] For the first five files, what is the chance that *one of them is highly complex*? **(D)**

**(23)** [4pts] For the first five files, what is the chance that *two or more are highly complex*? **(E)**

**(24)** [3pts] The simulation is not necessary to find the standard deviation of the sample mean. Using theory, when 20 fair dice are rolled, what is the s.d. of the sample mean? **(C)**

**(25)** [3pts] Why is there a discrepancy between the theoretical s.d. of the sample mean and the simulation results? **(A)**

**(26)** [4pts] In a roll of 20 dice, which of the following is the LEAST plausible result? **(C)**

**(27)** [4pts] If the Monte Carlo simulation is repeated but this time each sample has 100 observations (i.e. 100 dice rolled) instead of 20, what should you expect as the value of the 99<sup>th</sup> percentile in the new STATA summary of the simulation results for the *sample mean*? (Hint: $V[X] = \sigma^2 = 2.9167$ if $X$ records the value on a single rolled die.) **(E)**

**(28)** [4pts] If the Monte Carlo simulation is repeated but this time each sample has 100 observations (i.e. 100 dice rolled) instead of 20, which should you expect to INCREASE in the new STATA summary of the simulation results for the *sample median*? **(A)**

**(29)** [5pts] Suppose 60 percent of all unionized workers are in favor of a strike. If 132 unionized workers are randomly sampled and surveyed, what is the chance that less than half are in favor of a strike? **(A)**

**(30)** [5pts] Wait time for the "Airport Rocket" bus ($X$) is Uniformly distributed from 0 to 20 minutes. An airport monitor records wait times for 19 randomly selected travelers. (The chosen sampling procedure ensures that the independence assumption is reasonable.) What is the chance that the sample mean wait time is longer than 13 minutes? **(B)**

**(31)** [3pts] The CI estimator of a proportion is found with $\hat{P} \pm z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$. For a 97.5% CI, what is $z_{\alpha/2}$? **(A)**

**(32)** [4pts] Which of these would reduce the margin of error of the CI estimator of a proportion? **(B)**

**(33)** [2pts] What is the point estimate of the difference between the population proportions? **(B)**

**(34)** [4pts] With a 95% confidence level, what is the margin of error for the point estimate in the previous question? **(D)**

**(35)** [4pts] What does the 95% CI of the difference between proportions help make an inference about? **(D)**