1. **(Value 30%)** Suppose that the proportion of rural voters in a certain state who favor a particular gubernatorial candidate is 0.45 and the proportion of suburban and urban voters favoring the candidate is 0.60. If a sample of 200 rural voters and 300 urban and suburban voters is obtained, what is the approximate probability that at least 250 of these voters favor the candidate? Does the probability change substantially depending on whether you use a continuity correction in the approximation?

2. **(Value 25%)** The article “A Case Study of the Use of an Experimental Design in Preventing Shorts in Nickel-Cadmium Cells” (*J. Quality Technology*, 1988) states that at the beginning of the study about 6% of the cells produced were being scrapped because of internal shorts. In a sample of 235 cells produced under a particular set of trial conditions in the plant, there were 9 cells with internal shorts.

   (a) **(12%)** Do these data provide significant evidence that the new conditions have reduced the true proportion of manufactured cells with internal shorts? State the null hypothesis, the alternative hypothesis, the p-value for the test and your conclusion. State which output you used and why.

   (b) **(13%)** Provide a 95% confidence interval on the true proportion of cells with internal shorts if the new process is adopted. State the formula used to calculate this interval (use the more accurate formula from the text, not the simple formula). How does this interval compare to the interval produced by the `binom.test`? Are the intervals symmetric about the estimate $\hat{p}$?

   ```r
   > binom.test(9, 236, p = 0.06)
   Exact binomial test
   data: 9 and 236
   number of successes = 9, number of trials = 236, p-value = 0.2145
   alternative hypothesis: true probability of success is not equal to 0.06
   95 percent confidence interval: 0.01758380 0.07115644
   sample estimates: probability of success, 0.03813559
   > binom.test(9, 236, p = 0.06, alt = "less")
   number of successes = 9, number of trials = 236, p-value = 0.09503
   alternative hypothesis: true probability of success is less than 0.06
   95 percent confidence interval: 0.00000000 0.06560292
   > binom.test(9, 236, p = 0.06, alt = "greater")
   number of successes = 9, number of trials = 236, p-value = 0.948
   alternative hypothesis: true probability of success is greater than 0.06
   95 percent confidence interval: 0.02003644 1.00000000
   ```

3. **(Value 25%)** Extensive experience with fans of a certain type used in diesel engines has suggested the exponential distribution provides a good model for time until failure. Suppose that the mean time until failure is 25,000 hours. What is the probability that

   (a) **(15%)** A randomly selected fan will last at least 20,000 hours? At most 30,000 hours? Between 20,000 and 30,000 hours?

   (b) **(10%)** The lifetime of a fan exceeds the mean value by more than 2 standard deviations? More than 3 standard deviations?
4. (Value 20%) The article “Using Statistical Thinking to Solve Maintenance Problems” (*Quality Progress*, 1989) contains the 17 lifetimes, (10, 12, 15, 17, 18, 18, 20, 20, 21, 21, 23, 25, 27, 29, 29, 30, 35), measured as the number of 8-hour shifts that the roller lasted before needing to be replaced, of sinker rollers in a steel plant.

(a) (7%) Does the normal probability plot indicate that it is reasonable to assume these data come from a normal distribution? Indicate what pattern you wish to see in the plot and what patterns you do not want to see. Does this plot have the desired pattern?

(b) (6%) Provide a 95% confidence interval on the mean number of shifts that the sinker rollers last. State the formula used to calculate this interval.

(c) (7%) Provide a 95% prediction interval on the number of shifts that an as-yet unobserved sinker roller is expected to last. State the formula used to calculate this interval.

```r
> summary(lifetime); sd(lifetime) (= 6.731729)  
   Min.  1st Qu.   Median     Mean  3rd Qu.     Max.  
  10.000  18.000   21.000   21.760  27.000   35.000  
> t.test(lifetime)  
One Sample t-test  
data: lifetime  
t = 13.3306, df = 16, p-value = 4.421e-10  
alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:  
 18.30357 25.22584  
sample estimates:  
mean of x  
 21.76471  
> t.test(lifetime, mu = 1/0.05, alt = "less")  
t = 1.0809, df = 16, p-value = 0.8521  
alternative hypothesis: true mean is less than 20  
95 percent confidence interval: -Inf 24.61518  
> t.test(lifetime, mu = 1/0.05, alt = "greater")  
t = 1.0809, df = 16, p-value = 0.1479  
alternative hypothesis: true mean is greater than 20  
95 percent confidence interval: 18.91423 Inf
```