Last name: First name:

Student ID #: Section:

STAT 311 Midterm (**Show Your Work!**)

Fritz Scholz

1. **(7 points)** For a continuous random variable with uniform density over the interval $[40, 100]$ find the median and interquartile range.

   The interval can be split into 4 equal width (equal probability) adjacent subintervals $[40, 55), [55, 70), [70, 85), [85, 100]$ giving us median $= 70$ and IQR$= 85 - 55 = 30$.

2. **(8 points)** Suppose you are given the following ordered sample of size $n = 10$

   $$2.1,\ 2.4,\ 2.8,\ 3.2,\ 5.1,\ 6.5,\ 7.7,\ 10.2,\ 11.4,\ 12.0$$

   Give the plug-in estimates of all three quartiles.

   The three estimates are 2.8, $(5.1 + 6.5)/2 = 5.8$, and 10.2.

3. **(7 points)** Which six informative quantities does the R command `summary(x)` provide about a sample vector `x`?

   minimum, 1st quartile, mean, median, 3rd quartile and maximum.

4. **(9 points)** Explain briefly the distinction between **sampled distribution** and **sampling distribution**.

   The sampled distribution provides the sample $X_1, \ldots, X_n$ and the sampling distribution refers to the distribution of any statistics (say $\bar{X}_n$) computed from the sample.

5. **(6 points)** If a random variable X has a $\mathcal{N}(\mu, \sigma^2)$ distribution, what is the distribution of $[\,(X - \mu)\,/\,\sigma\,]^2$?

   It is a chi-square distribution with one degree of freedom, since $Z = (X - \mu)\,/\,\sigma \sim \mathcal{N}(0, 1)$.

6. **(6 points)** Give the mean and **variance** of $\bar{X}_{10}$ when it is computed for an i.i.d. random sample $X_1, \ldots, X_{10}$ from some distribution with mean $\mu = 1$ and **standard deviation** $\sigma = .2$. The final answers should consist of two decimal numbers.

   $E\bar{X}_{10} = \mu = 1$ and $\text{var}\,\bar{X}_{10} = \sigma^2/10 = (0.2)^2/10 = .004$.

7. **(5 points)** If `pnorm(1.5)` returns `0.9331928`, what would `pnorm(−1.5)` return?

   `pnorm(-1.5)` would return `.0668072=1-0.9331928`.

8. **(12 points)** A task takes a random time X to complete, where $EX = 2$ minutes and $\text{var}\, X = \frac{1}{36}$ minutes^2. What is the approximate chance that a person will complete 9 such tasks (one after the other) within 17 to 19 minutes, assuming that all 9 task times can be considered as independent random variables with the same distribution as X. Express the answer using the appropriate R function with specific decimal number arguments. Give a good numerical estimate for the answer returned by this R expression.

   $Y = X_1 + \ldots + X_9$, then $EY = 9 \cdot 2 = 18$ and $\text{var}\, Y = 9 \cdot (1/36) = (1/2)^2$. Thus

$$\begin{aligned} P(17 \le Y \le 19) &= \texttt{pnorm}(19, 18, 0.5) - \texttt{pnorm}(17, 18, 0.5) \\ &= \texttt{pnorm}(2) - \texttt{pnorm}(-2) \approx 0.95 \end{aligned}$$

9. **(3 points)** What does "i.i.d." stand for?

   independent identically distributed

10. **(6 points)** If X and Y are independent with means $\mu_x = 20$ and $\mu_y = 25$ and standard deviations $\sigma_x = 1$ and $\sigma_y = 2$, what are the mean and standard deviation of $6X - 4Y$? (Hint: the answer should consist of two clean integers)

    $E(6X - 4Y) = 6EX - 4EY = 120 - 100 = 20$ and $\text{var}(6X - 4Y) = 36 \cdot 1 + 16 \cdot 4 = 100 = 10^2$, thus the standard deviation of $6X - 4Y$ is 10.

11. **(8 points)** Explain the trade-off between the probability of type I error and the probability of type II error in hypothesis testing.

    As one gets larger the other gets smaller. Increasing the rejection region $\mathcal{R}$ decreases the acceptance region $\mathcal{R}^c$ and vice versa.

12. **(8 points)** Should a significance probability of .99 cause us to reject the null hypothesis H_0 when testing H_0 against some alternative H_1 at level $\alpha = 0.05$? Explain why or why not.

    Since the signifance probability 0.99 exceeds $\alpha = 0.05$ we cannot reject H_0 at this α.

13. **(5 points)** Fill in the question mark so that `pnorm(?)` produces the same as `pnorm(3,mean=4,sd=1)`.

    $? = (3 - 4)/1 = -1$, the standardized value.

14. **(10 points)** What command in R would give you $P(X \le 15)$ when X denotes the number of times that you get a face value of 5 or 6 in 60 independent rolls of a fair die?

    $X \sim \text{Binomial}(60, 1/3)$. Thus $P(X \le 15) = \texttt{pbinom}(15, 60, 1/3)$.