Problem 1: Discrete Probability Distribution and Expected Value

You are watching two items posted for sale on eBay and bid $30 for the first and $20 for the second item. You estimate that you are going to win the first bid with probability 0.1 and the second bid with probability 0.2, and the probability of winning both 0.04. Let X denote the random variable denoting the total amount of money you will spend on the two items.

(a) **List the sample space of all possible outcomes of winning or losing the two bids (there are 4 possible outcomes total). Use W for winning, L for losing. For instance, W_1L_2 represents winning the first bid but losing the second bid.**

$$S = \{W_1W_2, W_1L_2, L_1W_2, L_1L_2\}$$

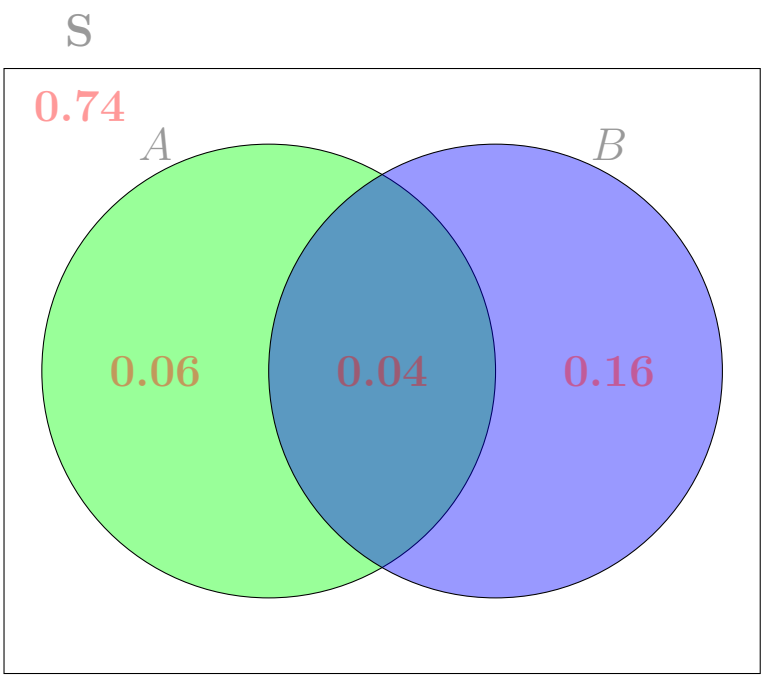
(b) **Find the probability of each outcome in the sample space.**

Let A = winning the 1st bid, B = winning the 2nd bid.

$$\begin{aligned}
P(W_1W_2) &= P(A \cap B) = 0.04 \\
P(W_1L_2) &= P(A \cap B^c) = P(A) - P(A \cap B) = 0.1 - 0.04 = 0.06 \quad \text{(by the partition rule)} \\
P(L_1W_2) &= P(A^c \cap B) = P(B) - P(A \cap B) = 0.2 - 0.04 = 0.16 \\
P(L_1L_2) &= P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - (0.1 + 0.2 - 0.04) = 0.74
\end{aligned}$$

Recall the information given in the question:

- Probability of winning BOTH bids is 0.04
- Probability of winning the first bid with probability 0.1
- Probability of winning the second bid with probability 0.2



(c) **For each outcome in the sample space, identify the total amount of money you will spend.**

$$X = \{0, 20, 30, 50\}$$

(d) **Suppose you could repeatedly bid on the two items over and over. Let X be the random variable that represents the total amount you will spend after each set of two bids. Based on your answer from part b) and c) (all possible values you could spend and its associated probability), find the probability distribution of X.**

- Winning the first bid costs \$30.
- Winning the second bid costs \$20.
- Losing both bids costs \$0.
- Winning both bids costs \$50.

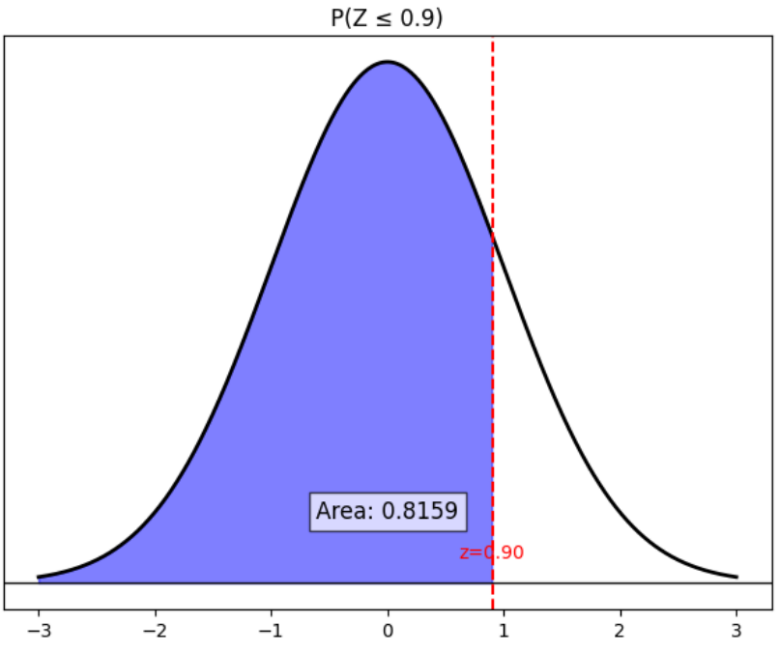
$$P(X = 0) = 0.74, \qquad P(X = 20) = 0.16,$$
$$P(X = 30) = 0.06, \qquad P(X = 50) = 0.04.$$

(e) **Modify the R command below to find the expected value of X.**

```
x <- c(0, 20, 30, 50)
p <- c(0.74, 0.16, 0.06, 0.04)
sum(x * p)
# Output: 7
```

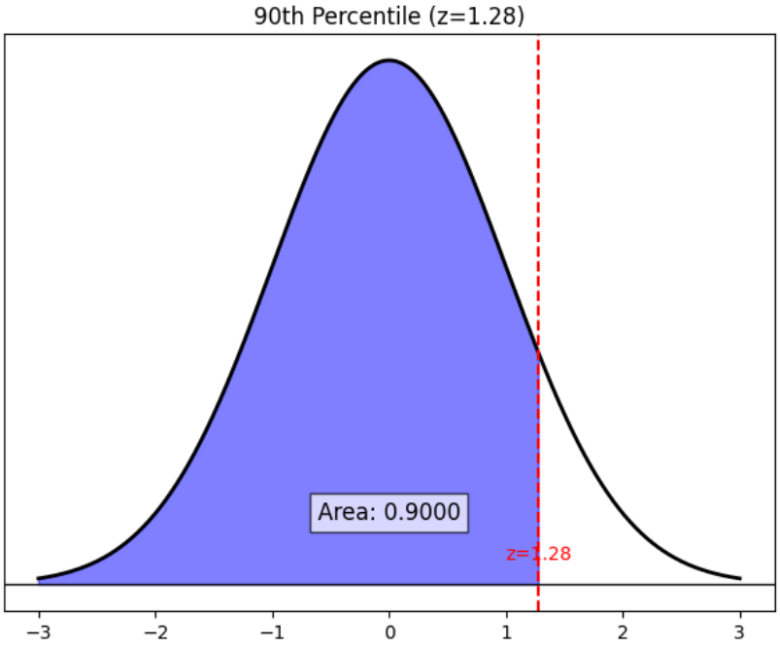
Problem 2: Standard Normal Distribution

(a) **Find** $P(Z < 0.9)$**.**



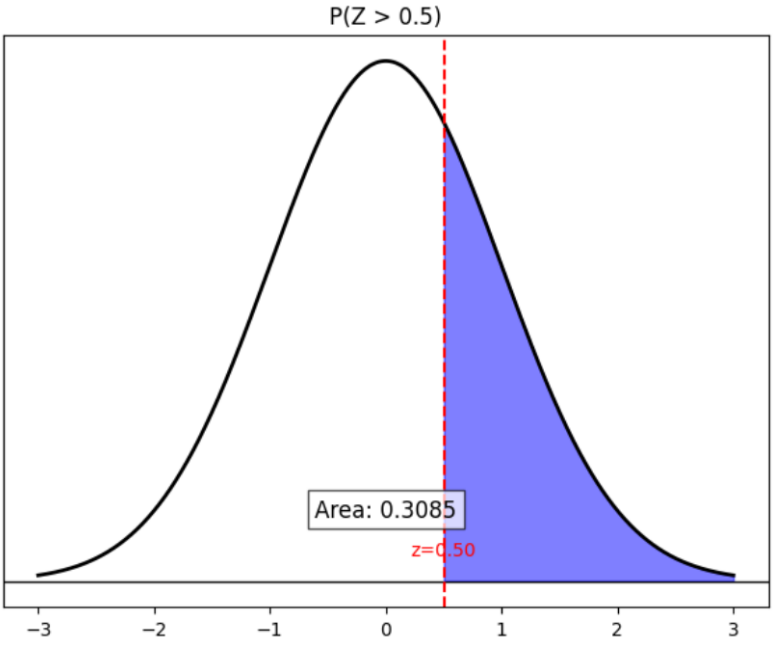

```
pnorm(0.9)

# Output: 0.8159
```

(b) **Find the z-value that marks the 90th percentile of the standard normal distribution.**



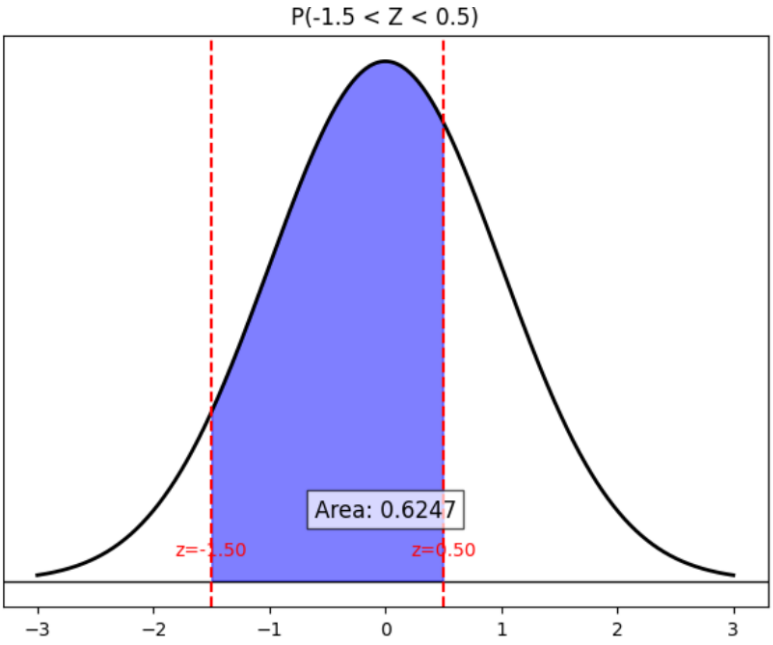

```
qnorm(0.9)
# Output: 1.2816
```

(c) **Find** $P(Z > 0.5)$**.**




```
1 - pnorm(0.5)
# Output: 0.3085
```

(d) **Find** $P(-1.5 < Z < 0.5)$**.**




```
pnorm(0.5) - pnorm(-1.5)
# Output: 0.6247
```

Problem 3: Use pnorm() command to find what percentage of the distribution is within 1 standard deviation from the mean in the standard normal distribution? Is it approximately 68%?

```
 pnorm(1)-pnorm(-1)
# Output: 0.6826895
```

Problem 4: Importing and Analyzing Data in R

1. **Import data set. Download survey_S25.txt (or csv) data from Canvas - Files - dataFile, then import it into R studio.**

```
survey <- read.csv("survey_S25.csv")
survey <- read.delim("survey_S25.txt")
```

2. **Examine the data set. The following commands take data as its argument.**

```
names(survey) # Column names
head(survey)  # First six rows
str(survey)   # Structure of dataset
```

3. **Create a frequency table of the variable "theater" (number of movies watched in the theater).**

```
table(survey$theater) / 289
round(table(survey$theater) / 289, 3)
```

4. **Convert "None" and "4 or more" to numbers and create a quantitative vector "theater.num" where "None" is 0 and "4 or more" is 4.**

```
theater.num <- ifelse(survey$theater == "None", 0,
                      ifelse(survey$theater == "4 or more", 4, survey$theater))
theater.num <- as.integer(theater.num)
```