Assume a discrete random variable X follows a discrete distribution, where the possible values of X are 0, 2, 4, 6, with the corresponding probabilities 0.3, 0.2, 0.2, and 0.3, respectively. Generate a sample of 10000 values from the given discrete distribution and see how close are the sample mean and standard deviation to the theoretical values

```
> # Define the distribution
> x <- c(0, 2, 4, 6)
> p <- c(0.3, 0.2, 0.2, 0.3)
>
> # Generate a sample of 10000 values
> set.seed(123) # set seed for reproducibility
> sample_x <- sample(x, size = 10000, replace = TRUE, prob = p)
>
> # Calculate sample mean and standard deviation
> sample_mean <- mean(sample_x)
> sample_sd <- sd(sample_x)
>
> # Calculate theoretical mean and standard deviation
> mu <- sum(x * p)
> sigma <- sqrt(sum(x^2 * p) - mu^2)
>
> # Print results
> cat("Sample mean:", sample_mean, "\n")
Sample mean: 3.0124
> cat("Sample SD:", sample_sd, "\n")
Sample SD: 2.416202
> cat("Theoretical mean:", mu, "\n")
Theoretical mean: 3
> cat("Theoretical SD:", sigma, "\n")
Theoretical SD: 2.408319
> |
```

From the observed data, it is evident that the computed sample mean and standard deviation are remarkably similar to the theoretical mean and standard deviation. This suggests that the 10,000-value sample effectively represents the specified discrete distribution.

Let random variable Y follows a binomial distribution, Binomial (n = 8, p = 0.5).

a). Calculate $P(Y = 5)$

$$P(Y = 5) = (8\ choose\ 5) * 0.5\text{^}5 * (1 - 0.5)\text{^}(8 - 5) = 0.21875$$

```
> p_y5 <- dbinom(5, size = 8, prob = 0.5)
> p_y5  # 0.21875
[1] 0.21875
> |
```

b) Calculate $P(Y < 3)$ and $P(Y > 6)$.

$$P(Y < 3) = P(Y = 0) + P(Y = 1) + P(Y = 2)$$
$$= (8 \, choose \, 0) * 0.5\verb|^|0 * (1 - 0.5)\verb|^|(8 - 0) + (8 \, choose \, 1) * 0.5\verb|^|1 * (1 - 0.5)\verb|^|(8 - 1) + (8 \, choose \, 2) * 0.5\verb|^|2 * (1 - 0.5)\verb|^|(8 - 2) = 0.109375$$

$$P(Y > 6) = P(Y = 7) + P(Y = 8)$$
$$= (8 \, choose \, 7) * 0.5\verb|^|7 * (1 - 0.5)\verb|^|(8 - 7) + (8 \, choose \, 8) * 0.5\verb|^|8 * (1 - 0.5)\verb|^|(8 - 8) = 0.0546875$$

```
> p_ylt3 <- pbinom(2, size = 8, prob = 0.5)
> p_ygt6 <- 1 - pbinom(6, size = 8, prob = 0.5)
> p_ylt3  # 0.109375
[1] 0.1445313
> p_ygt6  # 0.0546875
[1] 0.03515625
```

c) Draw 50,000 samples from this binomial distribution and draw a bar plot.

```
> set.seed(123)  # set seed for reproducibility
> samples <- rbinom(50000, size = 8, prob = 0.5)
> barplot(table(samples), main = "Binomial Distribution", xlab = "Y", ylab = "Frequency")
> 
```