

1 Test-style question on Transformations

See Chapter 5 of the course notes for lots of similar examples. You can use the formal transformation method, or just try from first principles. For example, the CDF for Y would be:

$$\begin{aligned}
 P(Y \leq y) &= P(X^2 \leq y) \\
 &= P(X \leq \sqrt{y}) \\
 &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\
 &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\
 &= F_X(\sqrt{y}) - (1 - F_X(\sqrt{y})) \\
 &= 2F_X(\sqrt{y}) - 1.
 \end{aligned}$$

Then to find the PDF for Y , we differentiate, using the chain rule:

$$\begin{aligned}
 f_Y(y) &= \frac{d}{dy} 2F_X(\sqrt{y}) \\
 &= 2f_X(\sqrt{y}) \times \frac{1}{2}y^{-1/2}.
 \end{aligned}$$

Since we are told that $X \sim N(0, 1)$, we know that

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Thus, subbing in $x = \sqrt{y}$ gives

$$\begin{aligned}
 f_Y(y) &= 2 \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2}y^{-1/2} \\
 &= \frac{1}{\sqrt{y}\sqrt{2\pi}} e^{-y/2} \\
 &= \frac{1}{\sqrt{2\pi y}} e^{-y/2}, \quad 0 \leq y < \infty.
 \end{aligned}$$

Thus, the correct answer is option **D**.

For the Canvas version of the question, full solutions, similar to those given above, would be required in order to gain full marks (10). For the Numbas version, just the correct choice would attract full marks (5).

2 Test-style question on bivariate distributions

See Chapter 6 of the course notes for similar examples – in particular, Section 6.1.

You should know that if $Z \sim \text{Exp}(\lambda)$, then

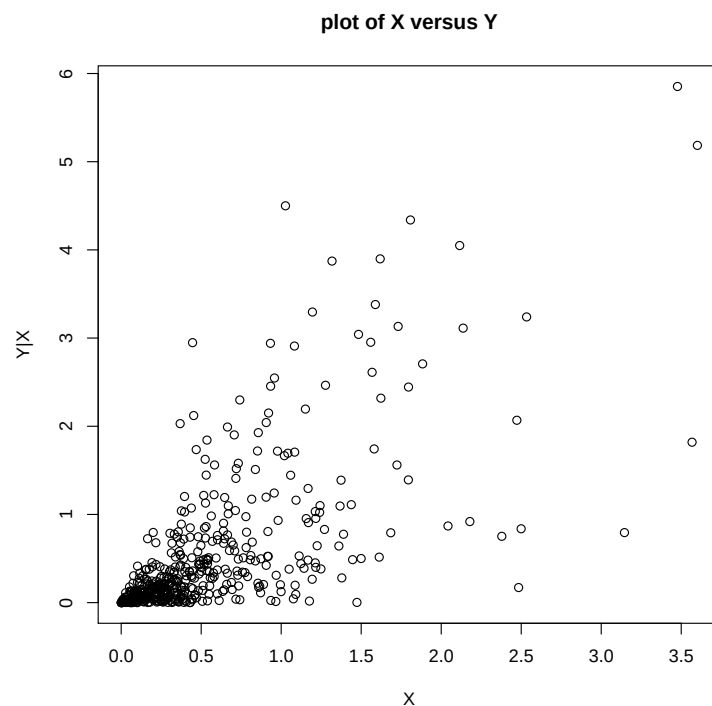
$$f_Z(z) = \lambda e^{-\lambda z}, \quad \lambda > 0.$$

Thus, it can be seen that $Y|X \sim \text{Exp}(1/X)$. So our R code might look something like:

```
1 bivariate = function(N){  
2   output = matrix(nrow = 2, ncol = N)  
3   X = rexp(N, 2)  
4   Y = rexp(N, 1/X)  
5   output[1,] = X  
6   output[2,] = Y  
7   return(output)}
```

Then we can put the function to good use – for example:

```
1 test = bivariate(500)  
2 plot(test[1,], test[2,])  
3 cor(test[1,], test[2,])  
4 [1] 0.67834786
```



For the Numbas version of this question, you would simply look for the plot in Numbas that most closely resembles the one you have produced (it will be obvious); you would then enter your correlation coefficient correct to 4 decimal places. For the Canvas version, you might copy-and-paste the `bivariate` function into a MS WORD document, along with the scatterplot produced in RStudio; you would also report your correlation coefficient.

3 Test-style question on bivariate distributions

Property 4 on page 51 of the course notes tells us that Z will be normally distributed, with mean and variance

$$\begin{aligned}\mu_z &= a\mu_x + b\mu_y \quad \text{and} \\ \sigma_z^2 &= a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\text{cov}(x, y),\end{aligned}$$

with

$$\mu_x = 4, \quad \mu_y = 1, \quad \sigma_x^2 = 8, \quad \sigma_y^2 = 5, \quad \text{cov}(x, y) = 2, \quad a = 1 \quad \text{and} \quad b = 3,$$

in this example. Thus, after substitution, we find that $Z \sim N(7, 65)$.

To produce a bivariate sample from (X, Y) , we can either use property 2 on page 51 to derive the conditional of $Y|X$ (or $X|Y$), and then write similar code to that in question 2; or we could use the `mvrnorm` function, as we do in Example 6.4 of the course notes. That is what we do here (less work involved!).

```
1 library(MASS)
2 npts = 1000
3 mu = c(4, 1)
4 sigma = matrix(data=c(8,2,2,5), nrow=2, byrow=TRUE)
5 test = mvrnorm(npts, mu, sigma)
```

Then we can easily find $P(Z > 8)$ and the correlation between X and Y :

```
1 X = test[,1]
2 Y = test[,2]
3 Z = X+3*Y
4 length(Z[Z>8])/npts
5 [1] 0.446
6 cor(X,Y)
7 [1] 0.2828786
```

For the Canvas version of this question, we require additionally the true values for $P(Z > 8)$ and the correlation. Since $Z \sim N(7, 65)$, we can work out $P(Z > 8)$ using tables, or by using the built-in function `pnorm`:

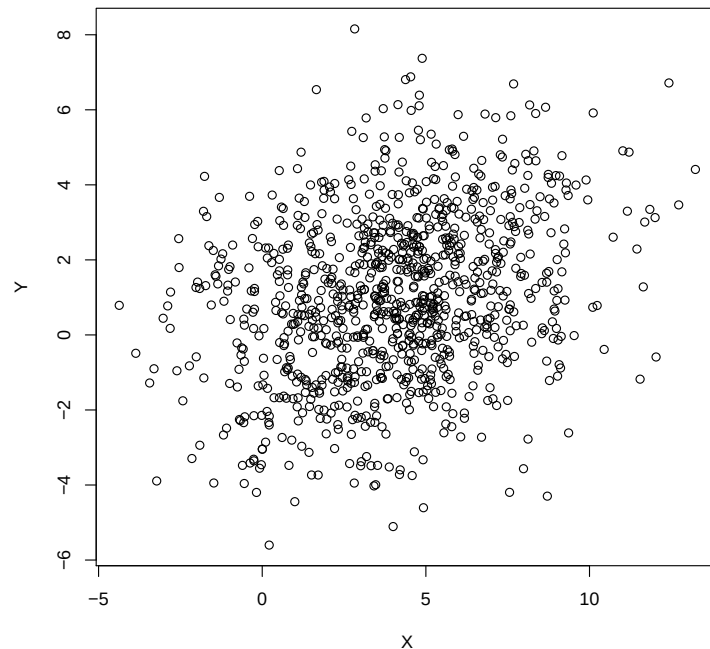
```
1 1 - pnorm(8, 7, sqrt(65))
2 [1] 0.4506439
```

For the true correlation between X and Y , we know that

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{2}{\sqrt{8 \times 5}} = 0.3162278.$$

For the Numbas version of the question, we would also need to produce a plot of X against Y , and compare to those provided:

```
1 plot(X, Y)
```



4 Test-style question on the Inverse CDF method

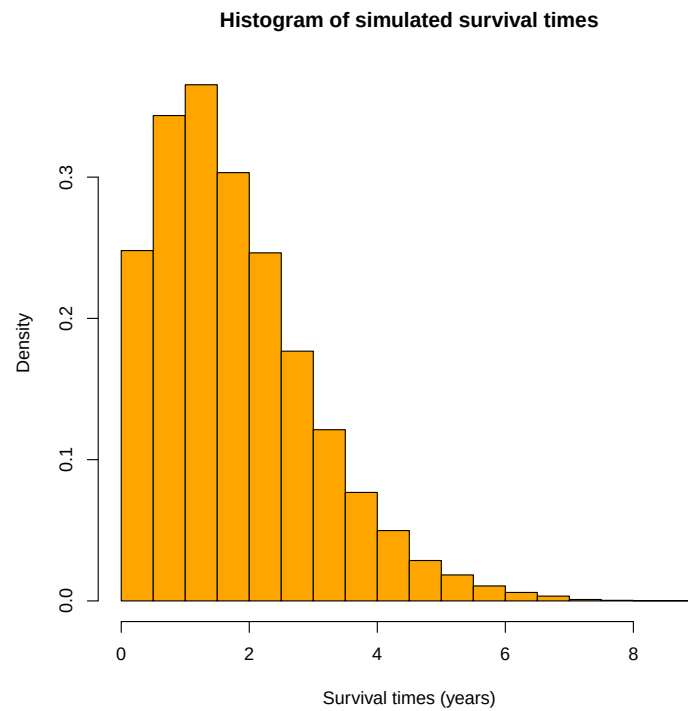
See Chapter 4 of the course notes for lots of similar examples.

First of all, we need to find the inverse of the CDF (note that if you were given the PDF, you would first need to integrate to find the CDF...). Doing so, you should find that

$$F_X^{-1}(x) = -\lambda [\log(1 - x)]^{1/\kappa},$$

and so we can now write our function to perform the inverse CDF method:

```
1 inv.cdf = function(N, lambda, kappa){
2   U = runif(N)
3   answer = lambda*((-log(1-U))^(1/kappa))
4   return(answer)}
5 test = inv.cdf(10000, 2, 1.5)
6 hist(test, col="orange", xlab="Survival times (years)", ylab="Density", freq=
  FALSE, main="Histogram of simulated survival times")
```



To find the probability of a patient surviving beyond 5 years:

```
1 length(test[test > 5]) / 10000
2 [1] 0.0201
```