CS 70 Fall 2020 — Discrete Mathematics and Probability Theory — HW 14

1 Short Answer

(a) Let X be uniform on the interval $[0,2]$, and define $Y = 2X+1$. Find the PDF, CDF, expectation, and variance of Y.

(b) Let X and Y have joint distribution

$$f(x,y) = \begin{cases} cxy + 1/4 & x \in [1,2] \text{ and } y \in [0,2] \\ 0 & \text{else} \end{cases}$$

Find the constant c. Are X and Y independent?

Solution:

(a) Let's begin with the CDF. It will first be useful to recall that

$$F_X(t) = \mathbb{P}(X \le t) = \begin{cases} 0 & t \le 0 \\ \frac{t}{2} & t \in [0,2] \\ 1 & t \ge 1 \end{cases}.$$

Since Y is defined in terms of X, we can compute that

$$\begin{aligned} F_Y[t] = \mathbb{P}(Y \le t) &= \mathbb{P}[2X+1 \le t] \\ &= \mathbb{P}\left[X \le \frac{t-1}{2}\right] \\ &= F_X\left(\frac{t-1}{2}\right) \\ &= \begin{cases} 0 & t \le 1 \\ \frac{t-1}{4} & t \in [1,5] \\ 1 & t \ge 5 \end{cases} \end{aligned}$$

where in the third line we have used the PDF for X. We know that the PDF can be found by taking the derivative of the CDF, so

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \begin{cases} \frac{1}{4} & t \in [1,5] \\ 0 & \text{else} \end{cases}. \tag{1}$$

By linearity of expectation $\mathbb{E}[Y] = \mathbb{E}[2X+1] = 2\,\mathbb{E}[X]+1 = 3$, and similarly

$$\text{Var}(Y) = \text{Var}(2X+1) = 4\,\text{Var}(X) = 4\frac{4}{12} = \frac{4}{3}$$

(b) To find the correct constant, we use the fact that a PDF must integrate to one. In particular,

$$1 = \int_1^2 \int_0^2 (cxy + 1/4)\, dy\, dx = 3c + 1/2,$$

so $c = 1/6$. In order to check independence, we need to first find the marginal distributions of X and Y:

$$f_X(x) = \int_0^2 f(x,y)\, dy = 1/2 + x/3$$
$$f_Y(y) = \int_1^2 f(x,y)\, dx = 1/4 + y/4.$$

Since $f_X(x) f_Y(y) = 1/8 + y/8 + x/12 + xy/12 \neq 1/4 + xy/6 = f(x,y)$, the random variables are not independent.

2 Continuous Probability Continued

For the following questions, please briefly justify your answers or show your work.

(a) Assume $\mathrm{Bob}_1, \mathrm{Bob}_2, \ldots, \mathrm{Bob}_k$ each hold a fair coin whose two sides show numbers instead of heads and tails, with the numbers on Bob_i's coin being i and $-i$. Each Bob tosses their coin n times and sums up the numbers he sees; let's call this number X_i. For large n, what is the distribution of $(X_1 + \cdots + X_k)/\sqrt{n}$ approximately equal to?

(b) If $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables of mean μ and variance σ^2, what is $\lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^n \frac{X_k - \mu}{\sigma n^\alpha} \in [-1,1]\right]$ for $\alpha \in [0,1]$ (your answer may depend on α and Φ, the CDF of a $N(0,1)$ variable)?

Solution:

(a) $\boxed{N\left(0, \sum_{i=1}^k i^2\right)}$.

$(X_1 + \cdots + X_k)/\sqrt{n} = \frac{X_1}{\sqrt{n}} + \cdots + \frac{X_k}{\sqrt{n}}$, and since each $\frac{X_i}{\sqrt{n}}$ converges to $N\left(0, i^2\right)$ by the central limit theorem, their sum must converge to $N\left(0, \sum_{i=1}^k i^2\right)$. Alternatively, if we let X_j^i be the j^{th} coin toss of Bob_i, then $(X_1 + \cdots + X_k)/\sqrt{n} = \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j^1 + \cdots + X_j^k)$. But the $Y_j = X_j^1 + \ldots X_j^k$ themselves are i.i.d. variables of mean 0 and variance $\sum_{i=1}^k i^2$, and so the central limit theorem again implies a limiting distribution of $N\left(0, \sum_{i=1}^k i^2\right)$ (this constitues an alternative proof of the fact that the sum of Gaussians is also a Gaussian, which we showed in class).

(b) $$\boxed{\lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^n \frac{X_k - \mu}{\sigma n^\alpha} \in [-1,1]\right] = \begin{cases} 1, & \text{if } \alpha > \frac{1}{2}, \\ \Phi(1) - \Phi(-1), & \text{if } \alpha = \frac{1}{2}, \\ 0, & \text{if } \alpha < \frac{1}{2} \end{cases}}.$$

For $\alpha > \frac{1}{2}$, the reasoning is exactly as in the law of large numbers: By Chebyshev's inequality, we have $1 - \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma n^{\alpha}} \in [-1,1]\right] = \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma n^{\alpha}} \notin [-1,1]\right] \leq \frac{1}{n^{2\alpha-1}} \xrightarrow{n\to\infty} 0$. The $\alpha = \frac{1}{2}$ case is a direct consequence of the central limit theorem, while the $\alpha < \frac{1}{2}$ case follows indirectly from it: $\mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma n^{\alpha}} \in [-1,1]\right] = \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma\sqrt{n}} \in \left[-\frac{1}{n^{\frac{1}{2}-\alpha}}, \frac{1}{n^{\frac{1}{2}-\alpha}}\right]\right]$
$\approx \mathbb{P}\left[N(0,1) \in \left[-\frac{1}{n^{\frac{1}{2}-\alpha}}, \frac{1}{n^{\frac{1}{2}-\alpha}}\right]\right] \xrightarrow{n\to\infty} 0.$

3 Exponential Distributions: Lightbulbs

A brand new lightbulb has just been installed in our classroom, and you know the life span of a lightbulb is exponentially distributed with a mean of 50 days.

(a) Suppose an electrician is scheduled to check on the lightbulb in 30 days and replace it if it is broken. What is the probability that the electrician will find the bulb broken?

(b) Suppose the electrician finds the bulb broken and replaces it with a new one. What is the probability that the new bulb will last at least 30 days?

(c) Suppose the electrician finds the bulb in working condition and leaves. What is the probability that the bulb will last at least another 30 days?

Solution:

(a) Let $X \sim \text{Exponential}(1/50)$ be the time until the bulb is broken. For an exponential random variable with parameter λ, the density function is $f_X(x) = \lambda\,\mathrm{e}^{-\lambda x}$ for $x > 0$. So in this case $\lambda = 1/50$. Thus we can integrate the density to find the probability that the lightbulb broke in the first 30 days:

$$\mathbb{P}[X < 30] = \int_0^{30} \left(\frac{1}{50} \cdot \mathrm{e}^{-x/50}\right) \mathrm{d}x = 1 - \mathrm{e}^{-30/50} = 1 - \mathrm{e}^{-3/5} \approx 0.451.$$

(b) The new bulb's waiting time Y is i.i.d. with the old bulb's. So the answer is

$$\mathbb{P}[Y > 30] = 1 - \mathbb{P}[Y < 30] = 1 - (1 - \mathrm{e}^{-3/5}) = \mathrm{e}^{-3/5} \approx 0.549.$$

(c) The bulb is memoryless, so the probability it will last 60 days given that it has lasted 30 days, is just the probability it will last 30 days:

$$\mathbb{P}[X > 60 \mid X > 30] = \mathbb{P}[X - 30 > 30 \mid X > 30] = \mathbb{P}[X > 30] = \mathrm{e}^{-3/5} \approx 0.549.$$

4 Useful Uniforms

Let X be a continuous random variable whose image is all of $\mathbb{R}$; that is, $\mathbb{P}[X \in (a,b)] > 0$ for all $a,b \in \mathbb{R}$ and $a \neq b$.

(a) Give an example of a distribution that X could have, and one that it could not.

(b) Show that the CDF F of X is strictly increasing. That is, $F(x+\varepsilon) > F(x)$ for any $\varepsilon > 0$. Argue why this implies that $F : \mathbb{R} \to (0,1)$ must be invertible.

(c) Let U be a uniform random variable on $(0,1)$. What is the distribution of $F^{-1}(U)$?

(d) Your work in part (c) shows that in order to sample X, it is enough to be able to sample U. If X was a discrete random variable instead, taking finitely many values, can we still use U to sample X?

Solution:

(a) Any random variable with density $f(x) > 0$ for all x works as a positive example; e.g. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (corresponding to the normal distribution) or $f(x) = \begin{cases} 1/2, & \text{if } |x| < 1, \\ \frac{1}{4|x|^2}, & \text{if } |x| \geq 1 \end{cases}$.

Any distribution of density f such that $f(x) = 0$ for all $x \in (a,b)$ for some $a,b \in \mathbb{R}, a \neq b$ works as a negative example; e.g. $f(x) = \begin{cases} e^{-x}, & \text{if } x \geq 0, \\ 0, & \text{otherwise} \end{cases}$ (corresponding to an exponential random variable) or $f(x) = \begin{cases} 1, & \text{if } x \in [0,1], \\ 0, & \text{otherwise} \end{cases}$ (corresponding to a uniform variable on $[0,1]$).

(b) $F(x+\varepsilon) = \mathbb{P}[X \leq x+\varepsilon] = \mathbb{P}[X \leq x] + \mathbb{P}[X \in (x, x+\varepsilon]] \geq F(x) + \mathbb{P}[X \in (x, x+\varepsilon)] > F(x)$, where in the very last inequality we used the fact that $\mathbb{P}[X \in (a,b)] > 0$ with $a = x$ and $b = x+\varepsilon$. To show invertibility, we need to show (i) injectivity and (ii) surjectivity. (i): If $x \neq y$, then either $x < y$ or $y < x$ and so either $F(x) < F(y)$ or $F(y) < F(x)$. In either case, $F(x) \neq F(y)$, and so F must be injective. (ii) F is continuous (in fact, differentiable with derivative f), approaching 1 as $x \to \infty$, and approaching 0 as $x \to -\infty$. Therefore, it must assume all values between 0 and 1, and hence is surjective.

(c) $\mathbb{P}\left[F^{-1}(U) \leq x\right] = \mathbb{P}[U \leq F(x)] = F(x)$, where $\{F^{-1}(U) \leq x\} = \{U \leq F(x)\}$ since F is strictly increasing. Thus $F^{-1}(U)$ and X have the very same CDF, which means that $F^{-1}(U)$ and X share the same distribution.

(d) Yes, we can! Assume X took values in a discrete set $\mathscr{A} = \{a_1, a_2, \ldots, a_n\} \subset \mathbb{R}$ with probabilities $\mathbb{P}[X = a_k] = p_k$. Then mimicking the argument from part (c), we can define $G : [0,1] \to \mathscr{A}$

as

$$G(x) = \begin{cases} a_1, & \text{if } x \le p_1, \\ a_2, & \text{if } x \in (p_1, p_1 + p_2], \\ a_3, & \text{if } x \in (p_1 + p_2, p_1 + p_2 + p_3], \\ \vdots & \vdots \\ a_{n-1}, & \text{if } x \in (\sum_{k=1}^{n-2} p_k, \sum_{k=1}^{n-1} p_k], \\ a_n, & \text{if } x \in (\sum_{k=1}^{n-1} p_k, 1] \end{cases}$$

(draw a picture of G's graph!), for which we have $\mathbb{P}[G(U) = a_k] = \sum_{j=1}^{k} p_j - \sum_{j=1}^{k-1} p_j = p_k = \mathbb{P}[X = a_k]$. That is, $G(U)$ and X have the same distribution as desired.

5 Uniform Means

To keep the doctor away, Bob goes to the supermarket to buy an apple. Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed uniform random variables on the interval $[0,1]$ (where n is a positive integer), where X_i is the quality of the ith apple Bob sees.

(a) Let $Y = \min\{X_1, X_2, \ldots, X_n\}$ be the quality of the worst apple Bob will see. Find $\mathbb{E}(Y)$. [*Hint*: Use the tail sum formula, which says the expected value of a nonnegative random variable is $\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > x)\,\mathrm{d}x$. Note that we can use the tail sum formula since $Y \ge 0$.]

(b) Let $Z = \max\{X_1, X_2, \ldots, X_n\}$ be the quality of the best apple Bob will see. Find $\mathbb{E}(Z)$. [*Hint*: Find the CDF.]

Solution:

(a) To calculate $\mathbb{P}(Y > y)$, where $y \in [0,1]$, this means that each X_i is greater than y, for $i = 1, \ldots, n$, so $\mathbb{P}(Y > y) = (1-y)^n$. We then use the tail sum formula:

$$\mathbb{E}(Y) = \int_0^1 \mathbb{P}(Y > y)\,\mathrm{d}y = \int_0^1 (1-y)^n\,\mathrm{d}y = -\frac{1}{n+1}(1-y)^{n+1}\Big|_0^1 = \frac{1}{n+1}.$$

Alternative Solution 1:

As explained above, $\mathbb{P}[Y \le y] = 1 - (1-y)^n$. This gives us the CDF, and if we take its derivative we'll get the probability density function $f(y) = n(1-y)^{n-1}$.

Then

$$\mathbb{E}(Y) = \int_0^1 y \cdot n(1-y)^{n-1}\,\mathrm{d}y.$$

Perform a u substitution, where $u = 1-y$ and $\mathrm{d}u = -\mathrm{d}y$. We see:

$$\mathbb{E}(Y) = n \cdot \int_0^1 -(1-u) \cdot u^{n-1}\,\mathrm{d}u = n \cdot \int_0^1 (u^n - u^{n-1})\,\mathrm{d}u = n\Big[\frac{u^{n+1}}{n+1} - \frac{u^n}{n}\Big]_{u=0}^{1}$$
$$= n\Big[\frac{(1-y)^{n+1}}{n+1} - \frac{(1-y)^n}{n}\Big]_{y=0}^{1} = n\Big[0 - \Big(\frac{1}{n+1} - \frac{1}{n}\Big)\Big] = n\Big[\frac{1}{n} - \frac{1}{n+1}\Big] = \frac{1}{n+1}.$$

Alternative Solution 2:
Consider adding another independent uniform variable X_{n+1}. $\mathbb{P}(X_{n+1} < Y)$ is just the probability that X_{n+1} is the minimum, which is $1/(n+1)$ by symmetry since all the X_i's are identical. It so happens that because X_{n+1} is a uniform variable on [0,1], this probability is equal to $\mathbb{E}(Y)$. Let f_Y denote the PDF of Y.

$$\begin{aligned}
\mathbb{P}(X_{n+1} < Y) &= \int_0^1 \mathbb{P}(X_{n+1} < y \mid Y = y) f_Y(y)\,\mathrm{d}y \\
&= \int_0^1 \mathbb{P}(X_{n+1} < y) f_Y(y)\,\mathrm{d}y && \text{(by independence)} \\
&= \int_0^1 y f_Y(y)\,\mathrm{d}y && \text{(CDF of the uniform distribution)} \\
&= \mathbb{E}(Y).
\end{aligned}$$

Alternative Solution 3:
Since $X_1,\ldots,X_n$ are i.i.d., their values split the interval $[0,1]$ into $n+1$ sections, and we expect these sections to be of equal length because they are uniformly distributed. Therefore, $\mathbb{E}(Y) = 1/(n+1)$, the position of the smallest indicator.

(b) We could use the tail sum formula, but it turns out that the CDF is in a form that makes it easy to take an integral. If $Z \le z$, where $z \in [0,1]$, each X_i must be less than z, which happens with probability z, so $\mathbb{P}[Z \le z] = z^n$. This gives us the CDF, and if we take its derivative we'll get the probability density function $f(z) = nz^{n-1}$. Then

$$\mathbb{E}(Z) = \int_0^1 z \cdot nz^{n-1}\,\mathrm{d}z = \int_0^1 nz^n\,\mathrm{d}z = \left[n \cdot \frac{z^{n+1}}{n+1}\right]_{z=0}^1 = \frac{n}{n+1}.$$

Alternative Solution:
As in the previous part, add another independent uniform random variable X_{n+1}. The probability $\mathbb{P}(X_{n+1} > Z)$ is just the probability that X_{n+1} is the maximum, which is $1/(n+1)$ by symmetry.

$$\begin{aligned}
\mathbb{P}(X_{n+1} > Z) &= \int_0^1 \mathbb{P}(X_{n+1} > z \mid Z = z) f_Z(z)\,\mathrm{d}z = \int_0^1 \mathbb{P}(X_{n+1} > z) f_Z(z)\,\mathrm{d}z \\
&= \int_0^1 (1-z) f_Z(z)\,\mathrm{d}z = \int_0^1 f_Z(z)\,\mathrm{d}z - \int_0^1 z f_Z(z)\,\mathrm{d}z \\
\frac{1}{n+1} &= 1 - \mathbb{E}(Z) \\
\mathbb{E}(Z) &= \frac{n}{n+1}
\end{aligned}$$

Alternative Solution 2:
Since $X_1,\ldots,X_n$ are i.i.d., their values split the interval $[0,1]$ into $n+1$ sections, and we expect these sections to be of equal length because they are uniformly distributed. The expectation of the smallest X_i is $1/(n+1)$, the expectation of the second smallest is $2/(n+1)$, etc. Therefore, $\mathbb{E}(Z) = n/(n+1)$, the position of the largest indicator.

Alternative Solution 3:
Let us define $Y_i = 1 - X_i$. Then, $Z = \max\{X_1, X_2, \dots, X_n\} = 1 - \min\{Y_1, Y_2, \dots, Y_n\}$. Observe that, although a function of X_i, the Y_i are also independent and identically distributed uniform random variables over $[0,1]$. Thus, we apply the previous part to find that $\mathbb{E}[\min\{Y_1, Y_2, \dots, Y_n\}] = 1/n+1$. As a result,

$$\begin{aligned}
\mathbb{E}[Z] &= \mathbb{E}[1 - \min\{Y_1, Y_2, \dots, Y_n\}] \\
&= 1 = \mathbb{E}[\min\{Y_1, Y_2, \dots, Y_n\}] \\
&= 1 - \frac{1}{n+1} \\
&= \frac{n}{n+1}
\end{aligned}$$

6 Darts but with ML

Suppose Alice and Bob are playing darts on a circular board with radius 1. When Alice throws a dart, the distance of the dart from the center is uniform $[0,1]$. When Bob throws the dart, the location of the dart is uniform over the whole board. Let X be the random variable corresponding to the distance of the player's dart from the center of the board.

(a) What is the pdf of X if Alice throws

(b) What is the pdf of X if Bob throws

(c) Suppose we let Alice throw the dart with probability p, and let Bob throw otherwise. What is the pdf of X (your answer should be in terms of p)?

(d) Using the same premise as in part c, suppose you observe a dart on the board but don't know who threw it. Let x be the dart's distance from the center. We would like to come up with a decision rule to determine whether Alice or Bob is more likely to have thrown the dart given your observation, x. Specifically, if we let A be the event that Alice threw the dart and B be the event that Bob threw, we want to guess A if $\mathbb{P}[A|X \in [x, x+dx]] > \mathbb{P}[B|X \in [x, x+dx]]$ (what do these two probabilities have to sum up to?). For what values of x would we guess A? (your answer should be in terms of p)

Solution:

(a) If Alice threw, then $X \sim U[0,1]$, so it's pdf is $f_{X|A}(x|A) = 1$. Note, the cdf is $\mathbb{P}[X < x|A] = \int_0^x 1dx = x$, which makes sense because this is exactly the area of a rectangle of length x and height 1.

(b) If Bob throws, then the probability that $X < x$ is equaled to the area of the disc of radius x around the center of the dartboard divided by the area of the dartboard. Thus, we have the cdf as:

$$\mathbb{P}[X < x|B] = \frac{\pi x^2}{\pi} = x^2$$
$$f_{X|B}(x|B) = \frac{d}{dx}\mathbb{P}[X < x|B] = 2x$$

(c) To find the pdf if X, we can again take the cdf first and take the derivative:

$$\begin{aligned}\mathbb{P}[X < x] &= \mathbb{P}[X < x|A]\mathbb{P}[A] + \mathbb{P}[X < x|B]\mathbb{P}[B] \\ &= px + (1-p)x^2 \\ f_X(x) &= p + 2(1-p)x\end{aligned}$$

(d) Intuitively, we can sketch out the pdfs of both Alice and Bob's throws and we see that Alice is more likely to hit closer to center compared to Bob. Thus it makes sense to say that there is a particular value x^* such that the distance of the dart from the center is less than x^*, then we guess Alice. Otherwise, we guess Bob. Specifically, we can compute with Bayes's rule:

$$\begin{aligned}\mathbb{P}[A|X \in [x, x+dx]] &= \frac{\mathbb{P}[X \in [x, x+dx]|A]\mathbb{P}[A]}{\mathbb{P}[X \in [x, x+dx]]} \\ &= \frac{f_{X|A}(x|A)dx * \mathbb{P}[A]}{f_X(X) * dx} \\ &= \frac{p}{p + 2(1-p)x}\end{aligned}$$

Note that this function is monotonically decreasing in x. In particular, we want to guess Alice if it is more likely that she threw the dart than Bob threw the dart, which means $\mathbb{P}[A|X \in [x, x+dx]] > 1/2$. Thus, we guess Alice if:

$$\begin{aligned}\frac{p}{p + 2(1-p)x} &> \frac{1}{2} \\ 2x(1-p) &< 2p - p \\ x &< \frac{p}{2(1-p)}\end{aligned}$$

Note that if we take $p = 1/2$ and plot out the conditional pdfs of Alice and Bob, we see that Alice's pdf is higher when $x < 1/2$ and Bob's pdf is higher when $x > 1/2$. Incidentally, if we take $p = 1/2$, we see that our decision boundary is exactly 1/2. Thus, the decision boundary corresponds to the point where the two pdfs, after scaling by p and $1-p$, have the same height.

7 Sampling a Gaussian With Uniform

In this question, we will see one way to generate a normal random variable if we have access to a random number generator that outputs numbers between 0 and 1 uniformly at random.

As a general comment, remember that showing two random variables have the same CDF or PDF is sufficient for showing that they have the same distribution.

(a) First, let us see how to generate an exponential random variable with a uniform random variable. Let $U_1 \sim Uniform(0,1)$. Prove that $-\ln U_1 \sim Expo(1)$.

(b) Let $N_1, N_2 \sim \mathcal{N}(0,1)$, where N_1 and N_2 are independent. Prove that $N_1^2 + N_2^2 \sim Expo(1/2)$.

Hint: You may use the fact that over a region R, if we convert to polar coordinates $(x,y) \to (r,\theta)$, then the double integral over the region R will be

$$\iint_R f(x,y)\,dx\,dy = \iint_R f(r\cos\theta, r\sin\theta) \cdot r\,dr\,d\theta.$$

(c) Suppose we have two uniform random variables, U_1 and U_2. How would you transform these two random variables into a normal random variable with mean 0 and variance 1?

Hint: What part (b) tells us is that the point (N_1, N_2) will have a distance from the origin that is distributed as the square root of an exponential distribution. Try to use U_1 to sample the radius, and then use U_2 to sample the angle.

Solution:

(a) The CDF of an exponential $Expo(\lambda)$ distribution is $1 - e^{-\lambda t}$. Let us prove that the $-\ln(U_1)$ also has the same CDF.

We see that

$$\begin{aligned} \mathbb{P}(-\ln(U_1) \le t) &= \mathbb{P}(\ln(U_1) \ge -t) \\ &= \mathbb{P}(U_1 \ge e^{-t}) \\ &= 1 - e^{-t} \end{aligned}$$

This shows that $-\ln(U_1)$ has an exponential distribution with $\lambda = 1$.

(b) We compute the CDF of $N_1^2 + N_2^2$. We want the probability that $N_1^2 + N_2^2 \le t$ for some t. This means that we are integrating the joint distribution over a circle of radius $\sqrt{t}$, centered at the origin. We therefore compute the following integral

$$\iint_{(x,y):x^2+y^2 \le t} \frac{1}{2\pi} e^{-(x^2+y^2)/2}\,dx\,dy = \int_0^{2\pi} \int_0^{\sqrt{t}} \frac{1}{2\pi} r e^{-r^2/2}\,dr\,d\theta$$

Evaluating this integral yields

$$\int_0^{2\pi} -\frac{e^{-r^2/2}}{2\pi}\Big|_0^{\sqrt{t}}\,d\theta = \int_0^{2\pi} \frac{1 - e^{-t/2}}{2\pi}\,d\theta = 1 - e^{-t/2}.$$

This proves that $N_1^2 + N_2^2 \sim Expo(1/2)$.

(c) We will sample the point (N_1, N_2) using uniform random variables U_1 and U_2. We first sample the radius, which we know is an exponential distribution. Therefore, we know that $-2\ln(U_1)$ is an exponential $1/2$ distribution, so $\sqrt{-2\ln(U_1)}$ can be our radius. Since the (N_1, N_2) joint distribution is rotationally symmetric, we know that we can pick our angle uniformly at random once the radius is determined. Therefore, we let $\theta = 2\pi U_2$.

We will actually arrive at two Gaussians, so we can just take N_1, which will be

$$\boxed{\sqrt{-2\ln(U_1)}\cos(2\pi U_2)}$$