Sample Problems

Math 394

1 Quizzes

1.1

Cross the circles corresponding to possible situations for a generic random variable X. Leave the circles corresponding to impossible situations blank

⊗ $EX = 1, EX^2 = 2$: $EX^2 = 2 > (EX)^2 = 1$, so all's right.

⊗ $EX = 2, EX^2 = 4$: this is borderline, since $EX^2 - (EX)^2 = Var[X] = 4 - 4 = 0$, i.e., this RV is actually the fixed constant 2.

○ $EX = 0, Var[X] = -1$: Variances are positive (or zero in degenerate cases like the previous question).

⊗ $EX = 2, Var[X] = 1$: besides the requirement $Var[X] \geq 0$, there is no special limitation on expected values and variances.

1.2

Let X be a continuous RV. Then it is always true that (f_X is the density, and F_X is the cdf of X)

○ $f_X(x) > 0 \quad \forall x \in \mathbb{R}$: it easily happens that $f_X(x) = 0$ over a large set of values

○ F_X is *strictly* increasing for all $x \in \mathbb{R}$: same as above (since $f_X = F'_X$ the two statements are *almost*[1] equivalent

[1]They are if we restrict our considerations to RVs with continuous densities (except at, possibly, a finite number of points). In the most general case, an increasing function like F might not be differentiable on a countable set of values - but we'd need more advances analysis to work this out.

⊗ F_X is differentiable for all $x \in \mathbb{R}$, except, possibly, at a countable number of values. That's true, even in more general cases than the ones we are considering (see footnote)

1.3

If $P[X \geq 0] = 1$, then

○ $Var[X] > 0$: you see, it could still be possible that X was a degenerate RV (i.e., a constant), with zero variance.

⊗ $EX \geq 0$: this follows easily from the definition.

○ $EX \geqslant \sqrt{Var[X]}$ There is no relationship between expectation and variance

2 Problems

2.1

Suppose X is a normal RV, with mean μ, and variance σ^2. Find a and b such that $aX + b$ is normal with mean 0, and variance $\frac{\sigma^2}{4}$.

Solution: $E[aX+b] = aEX + b = a\mu + b$, and $Var[aX+b] = a^2 Var[X] = a^2\sigma^2$. Hence, we need $a^2 = \frac{1}{4}$, and $\frac{\mu}{2} + b = 0$, or $b = -\frac{\mu}{2}$. The requested RV is $\frac{1}{2}X - \frac{\mu}{2} = \frac{X-\mu}{2}$.

2.2

Suppose X has a survival function of the form

$$R_X(t) = P[X > t] = \begin{cases} 1 & t \leq 0 \\ e^{-t^\alpha} & t > 0 \end{cases} \tag{1}$$

where $\alpha > 0$ is a constant. This family of distributions (as αvaries) is known as the family of *Weibull* distributions, and is very popular in survival analysis.

1. Write the density and cdf of X
2. Calculate $P[t + \varepsilon > X > t \,|X > t] = h_\varepsilon(t)^2$, and

$$h(t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} h_\varepsilon(t) \tag{2}$$

(this is called the "hazard function" for the distribution of X.

3. The hazard function is an indication of how likely it is for something to fail, given that it has survived until now, when X represents the time to failure of something. Consider the three cases $\alpha < 1, \alpha = 1, \alpha > 1$: what can you say in terms of failure risk, as time goes by?

Solution: The Weibull distribution is discussed on p. 220 of the book, and, more generally than in (1), is defined by

$$R(x) = e^{\beta x^{\alpha}}$$

for $x > 0$, and $R = 1$ for $x < 0$. The book uses a different form for the exponent, but, by a change of origin, it amounts to the same.

1. $F_X = 1 - R_X = 1 - e^{-t^{\alpha}}$ (for $t \geq 0$), and $F_X = 0$ for $t < 0$. The density is the derivative of F_X, hence

$$f_X(t) = \begin{cases} 0 & t < 0 \\ \alpha t^{\alpha-1} e^{-t^{\alpha}} & t > 0 \end{cases}$$

(you can define the value at $t = 0$ however you wish, since only integrals of f_X have real meaning)

2. As we have seen repeatedly in problems involving the exponential distribution, the routine goes like this:

$$P\left[X > t + \varepsilon \,|X > t\right] = \frac{P\left[t + \varepsilon > X > t \cap X > t\right]}{P\left[X > t\right]} =$$

$$= \frac{P\left[t + \varepsilon > X > t\right]}{P\left[X > t\right]} = 1 - \frac{e^{-(t+\varepsilon)^{\alpha}}}{e^{-t^{\alpha}}} =$$

$$= 1 - \exp\left\{t^{\alpha} - (t + \varepsilon)^{\alpha}\right\} = h_{\varepsilon}(t)$$

Deriving $-e^{-t^{\alpha}}$ with respect to t, and dividing by R (that's what taking the limit in (2) means), we get

$$h(t) = \alpha t^{\alpha-1}$$

Notice that with this definition, the relation between h and he other means of describing a distribution for a nonnegative RV is

$$R(t) = e^{-\int_0^t h(s)ds}$$

for $t > 0$, and $R = 1$ for $t < 0$.

3. $\alpha = 1$ corresponds to an exponential distribution with parameter 1 - the hazard function is constant, as there is no change in the probability of imminent failure, as time goes by, $\alpha > 1$ leads to an increasing hazard (t is raised to a positive power), and $\alpha < 1$, to a decreasing hazard (t is raised to a negative power).

2.3

Two pieces of machinery produce 1 product item per unit time. Let X and Y be their TTF (time to failure: the time lapse from startup to breakdown), and assume that they break independently, and that their "lifetimes" have a common cdf F. They can be set up either to work in parallel (hence, producing 2 items per unit time, as long as both are working, then 1, after the first breakdown, finally nothing), or on "stand-by" (i.e., one starting work as soon as the first breaks down - hence, from overall start, until the second breakdown, 1 unit per time is produced).

1. Taking note that

$$1 - F_{\min(X,Y)} = (1 - F_X)(1 - F_Y) \tag{3}$$

while

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dx \tag{4}$$

write a formula for (your choice: find the density, or the cdf, or the survival function) the overall production of the two setups

2. Write the explicit formulas, when the common distribution is the uniform distribution over $[0, 1]$

3. Write the explicit formulas, when the common distribution is exponential with parameter λ

4. [much harder] Prove relation (3)

5. [even worse] Prove relation (4)

Solutions:

1. Note that $\min\{X,Y\} > z$ if and only if both $X > z$ and $Y > z$, while $\max\{X,Y\} \le z$ if and only if both $X \le z$, and $Y \le z$. Also, since X and Y are independent,

$$P[X \geq> z, Y > z] = P[X > z] P[Y > z] = R_X(z) R_Y(z)$$

(this answers question #4)

$$P[X \le z, Y \le z] = P[X \le z] P[Y \le z] = F_X(z) F_Y(z)$$

In a "parallel" setup, they will produce 2 items per unit time between time 0, and time $\min\{X,Y\}$, and 1 item per unit time between time $\min\{X,Y\}$, and time $\max\{X,Y\}$. Total production is then

$$S = 2 \cdot \min\{X,Y\} + (\max\{X,Y\} - \min\{X,Y\}) = \min\{X,Y\} + \max\{X,Y\}$$

To compare with the "stand-by" setup, we can condition on whether X or Y is the minimum (or maximum). Since they are independent, and have the same distribution, the two possibilities have each probability $\frac{1}{2}$. Hence, conditioning on which is which, the collective production has distribution

$$P[S \leq t] = \frac{1}{2} P[X+Y \leq t | X < Y] + \frac{1}{2} P[Y+X \leq t | Y < X] =$$

$$\frac{1}{2}\left(\frac{P[X+Y \leq t, X<Y]}{P[X<Y]} + \frac{P[X+Y \leq t, Y<X]}{P[Y<X]}\right)$$

However, by symmetry, $P[X<Y] = P[Y<X] = \frac{1}{2}$, and we arrive at

$$P[X+Y \leq t, X<Y] + P[X+Y \leq t, Y<X] = P[X+Y \leq t]$$

In a "standby" setup, only one item is produced per unit time, between 0 and $X+Y$. The formula (4) gives us the density of this total work time, but, in generic form, this can be characterized by

$$P[S \leq t] = P[X+Y \leq t]$$

Hence, the two setups are equivalent.

2. If the two variables are uniform over $[0,1]$, the sum $X+Y$ has, by (4), density

$$\int_0^1 1_{[0,1]}(t-x) 1_{[0,1]}(x) dx = t 1_{[0 \leq t < 1]} + (2-t) 1_{[1 \leq t \leq 2]}$$

(cf. the book at p. 261). Hence,

$$F(t) = \begin{cases} \frac{t^2}{2} & 0 \leq t \leq 1 \\ 2t - 1 - \frac{t^2}{2} & 1 \leq t \leq 2 \end{cases}$$

3. The exponential case is special, in that the equivalence can be seen also by taking advantage of the peculiarities of this distribution. In fact, in this case (of course, $t > 0$), $\tau_1 = \min\{X, Y\}$ is distributed like an exponential with parameter $2 \cdot \lambda$ (since its survival function is the product of the two survival functions, $e^{-\lambda t} \cdot e^{-\lambda t} = e^{-2\lambda t}$). As soon as the first component fails, the second keeps going, but, since the exponential distribution "has no memory", it starts with a fresh survival function $e^{-\lambda t}$, independently of what happened until then. Thus, calling τ_2 the failure time, counted from the first failure, the total time of operation is $\tau_1 + \tau_2$, and the production is $2 \cdot \tau_1 + \tau_2$. But it is easy to check that $2 \cdot \tau_1$, twice an exponential RV with parameter 2λ, has exponential distribution with parameter λ: $P[2\tau_2 > t] = P\left[\tau_2 > \frac{t}{2}\right] = e^{-2\lambda \cdot \frac{t}{2}} = e^{-\lambda t}$! Hence, the total production has the distribution of the sum of two independent RVs with parameter λ. On the other hand, the "stand-by" operation has a total production of

$X+Y$, the sum of the two independent failure times, both exponential with parameter λ. In other words, in this particular case, both arrangements behave exactly the same. The distribution of this quantity, by (4), has density

$$f_{X+Y}(z) = \int_0^\infty \lambda^2 e^{-\lambda(z-y)} 1_{\{z>y\}} e^{-\lambda y} dy = \lambda^2 \int_0^z e^{-\lambda z} dy =$$

$$= \lambda^2 z e^{-\lambda z}$$

(this is known as a "Gamma distribution, with parameters 2 and λ - the exponential is a $\Gamma(1, \lambda)$, and the sum of n independent exponentials, each with parameter λhas a $\Gamma(n, \lambda)$ distribution: cf. sec. 5.6.1 and p. 262ff. of the book).

4. We already solved this problem in question #1
5. This a famous formula, and the combination of the two densities is known as the "convolution product" of the two (denoted by $f_X * f_Y$). You can look up the proof section 6.3 (it is formula 3.2) of the book (p. 260-261)

More Problems

3 Quizzes

3.1

Which of the following is a legitimate probability density? Cross the circles corresponding to the ones that are, leaving the others blank (we assume $\lambda > 0$ everywhere below)

○ $f(x) = \lambda e^{-\lambda x}$ for $x < 0$, 0 otherwise. $e^{-\lambda x} \to \infty$, as $x \to \infty$, hence f cannot be integrable.

○ $f(x) = 1$ for $-\infty < x < \infty$. Constants are not integrable over infinite intervals.

○ $f(x) = 2\lambda e^{-\lambda|x|}$ for all $x \in \mathbb{R}$. f is integrable, however

$$2\int_{-\infty}^{\infty} \lambda e^{-\lambda|x|} dx = 2 \cdot 2 \int_0^\infty \lambda e^{-\lambda x} dx = 4 \cdot 1$$

The constant in front should have been $\frac{1}{2}$, for this to be a density (that is sometimes called a "double-sided exponential").

○ $f(x) = 2\lambda e^{-\lambda x}$ for all $x \in \mathbb{R}$. Just like in the first example, this function is not integrable at $-\infty$.

⊗ $f(x) = \frac{1}{x}$ for $x \in [1, e]$, 0 everywhere else. $\frac{1}{x} > 0$ in the interval, and $\int_1^e \frac{dx}{x} = \log e - \log 1 = 1$

3.2

Let $f(x) = (1+\theta)\, x^\theta$ for $0 \le x \le 1$, and $f = 0$ otherwise. This is a probability density *if and only if*

○ $\theta \ge 0$. f is a density, but there are other values for which it is, namely $-1<\theta<0$

○ $\theta \ge 1$. Same as above.

⊗ $\theta>-1$ Clearly, $\int_0^1 (1+\theta)x^\theta\, dx=1$ in this case. Note that $\theta=-1$ does not work either, because the "density" would turn out to be identically zero.

4 Problems

4.1

Suppose that earthquakes register an intensity of M on the Richter scale, where, for $M > 3$, $M - 3$ has exponential distribution, with parameter $\lambda = 2$. We consider only these earthquakes.

1. Compute EM, and $Var[M]$
2. Write the density function, or the cdf, or the survival function for M (remember that we are ignoring the case $M \le 3$, so that, as far as we are concerned, $P[M \le 3] = 0$)
3. Assuming we observe two earthquakes, and that their intensities may be assumed to be independent, what is the probability that the *smallest* has intensity greater than 4?

Solution: Note that, since

$$P[M-3>m] = e^{-2m} \quad m>0, \text{ otherwise } = 1$$

we will have

$$P[M>m] = \begin{cases} 1 & m \le 3 \\ e^{-2\cdot(m-3)} & m > 3 \end{cases}$$

(since $M > m \Rightarrow M - 3 > m - 3$). This, essentially, solves question 2.

1. We can either note that $f_M(m) = -\frac{dR_M}{dm}$, or refer to a useful fact: for nonnegative RVs,

$$EX = \int_0^\infty R_X(x)dx$$

The proof is easy, if we integrate by parts, remembering that $R \to 0$ at $x \to \infty$ (in fact, for integrability, it is easy to check that we need $xR_X \to 0$), and that $f_X = -\frac{dR}{dx}$:

$$EX=\int_{-\infty}^{+\infty} x f_X(x)\,dx=\left[x R_X(x)\right]_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty} R_X(x)\,dx=0+\int_{-\infty}^{+\infty} R_X(x)\,dx$$

In our case,

$$EM = \int_0^3 dm + \int_3^\infty e^{-\lambda(m-3)} dm = 3 + \frac{1}{\lambda}$$

(after an easy substitution). Of course, it would be even faster to note that $M-3$ is exponential with parameter λ, hence $E[M-3] = EM - 3 = \frac{1}{\lambda}$. This way we can get the variance without even working: $Var[M-3] = Var[M] = \frac{1}{\lambda^2}$.

2. We computed R_M in the preamble. Taking the negative of the derivative, gives

$$f_M = \begin{cases} 0 & m < 3 \\ \lambda e^{-\lambda(m-3)} & m > 3 \end{cases}$$

and

$$F_M = 1 - R_M = \begin{cases} 0 & m < 3 \\ 1 - e^{-\lambda(m-3)} & m > 3 \end{cases}$$

3. Given two independent RVs, M_1, and M_2,

$$\min\{M_1, M_2\} > m \Leftrightarrow M_1 > m, M_2 > m$$

and

$$R_{\min\{M_1,M_2\}}(m) = P[M_1 > m, M_2 > m] = R_{M_1}(m) R_{M_2}(m)$$

Hence

$$P[\min\{M_1, M_2\} > 4] = e^{-\lambda} \cdot e^{-\lambda} = e^{-2\lambda}$$

4.2

We are observing raindrops falling on a specific spot. Assume that the number of raindrops over time $t-s$, $N_t - N_s$ $(0 \le s \le t)$, has a Poisson distribution of parameter $30(t-s)$ (time is measured in minutes), and that, for disjoint intervals $0 \le s \le t \le v \le u$, the RVs $N_u - N_v$, and $N_t - N_s$ are independent. This family of RVs $\{N_t\}$ is called a *Poisson Process*.

1. What is the probability of no raindrops falling between time 0 and time 10?
2. Assume no raindrops fell in the first 5 minutes. What is the *conditional* probability that no drops will fall in the following 10 minutes? Compare with your result in #1.

Solution: We have quickly mentioned in class how the times between arrivals in a Poisson process are independent exponentials. This points to one way of solving this problem (we are dealing with the first arrival time, which is an exponential with parameter 30). But we can also proceed directly from the Poisson distribution:

Solution: We have quickly mentioned in class how the times between arrivals in a Poisson process are independent exponentials. This points to one way of solving this problem (we are dealing with the first arrival time, which is an exponential with parameter 30). But we can also proceed directly from the Poisson distribution:

1. If no raindrops fall in the first 10 minutes, $N_{10} = 0$, and

$$P[N_{10} = 0] = \frac{(10\lambda)^0}{0!} e^{-10\lambda} = e^{-10\lambda} = e^{-300} \simeq 0$$

since N_{10} is a Poisson RV with parameter $10\lambda = 10 \cdot 30$. Note that $\lambda = 30$ means that, on average, we should observe 30 raindrops a minute, which explains the minuscule probability. Equivalently, calling T_1 the time when the first raindrop falls, the statement means $T_1 > 10$, and

$$P[T_1 > 10] = e^{-10\lambda} = e^{-300} \simeq 0$$

2. The answer is instantaneous, whichever approach we take. Whatever happened in the first 5 minutes, i.e. whatever the value of N_5, $N_{15} - N_5$ is independent of N_5, and is a Poisson with parameter $30 \cdot (15 - 5 = 300)$. Hence, the result is, again, e^{-300}. Equivalently, the defining property of exponential RVs is that

$$P[T_1 > t + s \,|T_1 > s] = P[T_1 > t]$$

so that we find, again, the same result.

4.3

Noting that a binomial RV with parameters n, p can be thought as the sum of n independent Bernoulli RVs with parameter p (it has the same distribution as such a sum), what will be the distribution of the sum of two independent binomial RVs, with respective parameters n, p and m, p?

Solution: Regardless of how the binomial RV with parameters n, p (call it Z) was constructed, it has the same distribution as as $\sum_{k=1}^{n} X_k$, where X_k are independent Bernoulli RVs with parameter p. Hence if Z_1 is $bin(n, p)$, and Z_2 is $bin(m, p)$, $Z_1 + Z_2$ will be distributed like

$$\sum_{k=1}^{n} X_k + \sum_{k=n+1}^{n+m} X_k = \sum_{k=1}^{n+m} X_k$$

where all the X_k are independent Bernoulli with parameter p. Hence the final distribution will be $bin(n + m, p)$.

4.4

Noting that, if X_k is a sequence of Bernoulli RVs, with parameter p, the distribution of $\sum_{k=1}^{n} \frac{X_k - p}{\sqrt{np(1-p)}}$ is approximately a Gaussian (Normal) distribution of parameters $\mu = 0$, and $\sigma^2 = 1$, while the sum $\sum_{k=1}^{n} X_k$ is a binomial with parameters n, p, using the result in problem #4.3, explain intuitively what you believe will be the distribution of the sum of two independent Normal RVs, with respective means μ_1, μ_2, and variances σ_1^2, σ_2^2. Remember, as useful information, that

- For two RVs, X, Y, $E[X+Y] = EX + EY$
- For two *independent* RVs, $Var[X+Y] = Var[X] + Var[Y]$
- If X is $N(\mu, \sigma^2)$, $\frac{X-\mu}{\sigma}$ is $N(0,1)$

Solution: It is hard to provide a handwaving argument that will look "obvious" to everyone (that's why we look for *proofs*: they are much more universally convincing). Anyway, let's say that, given that $Z_1 + Z_2$ will be a binomial RV with expectation the sum of the expectations, and variance equal to the sum of variances, we expect the limit (in the Central Limit) to be a Gaussian with the corresponding parameters. There should be no trouble interchanging the sum of two RVs with their limit, so we expect the sum of our two normals $N(\mu_i, \sigma_i^2)$ to be normal, $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2)$ (which it is!)

4.5

A reasonable model for shooting at a circular target, is to assume that the point that is hit will be at a distance (in some units) R from the center, such that R^2 is exponential with parameter λ (the value of λ would measure your marksmanship). Suppose a sharpshooter tries, with $\lambda = 4$.

1. If the target has radius 4, what is the probability of missing the target altogether?
2. What is the probability of scoring a bull's eye if, for this purpose, you have to hit within a radius of .1 from the center?
3. Using the observation that, for two independent exponential RVs, $R_{\min\{X,Y\}} = R_X R_Y$, if two sharpshooters, with the same $\lambda = 4$, shoot at the same target, what is the probability that the *winner* will have scored a bull's eye?

Solution: I will write a short proof of the following statement, but it relies on the theory of multiple integrals, so it is here just for the fun of it: *If X and Y are two independent RVs, both distributed as Gaussians with $\mu = 0$,* *and* $\sigma^2 = a^2$, *then* $X^2 + Y^2$ *is an exponential, with parameter* $\frac{1}{2a^2}$. The proof goes as follows

$$P\left[X^2 + Y^2 \leq r^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{X^2+Y^2\leq r^2}\frac{1}{\sqrt{2\pi a^2}}\frac{1}{\sqrt{2\pi a^2}}e^{-\frac{x^2}{2a^2}}e^{-\frac{y^2}{2a^2}}dxdy =$$

$$= \frac{1}{2\pi a^2}\int_0^r\int_0^{2\pi} e^{-\frac{\rho^2}{2a^2}}\rho d\rho d\theta$$

(where we have changed to polar coordinates, to take advantage of the symmetry). Integrating first in $d\theta$,

$$= \frac{1}{a^2}\int_0^r \rho e^{-\frac{\rho^2}{2a^2}}d\rho$$

and, with the substitution $x = \frac{\rho^2}{2a^2} \Rightarrow dx = \rho d\left(\frac{\rho}{a^2}\right)$,

$$= \int_0^{\frac{r^2}{a^2}} \exp\left\{-x\right\}dx = 1 - e^{-\frac{r^2}{2a^2}}$$

i.e., $X^2 + Y^2$ is an exponential RV with parameter $\frac{1}{2a^2}$. Incidentally, the same trick allows us to prove that $\frac{1}{\sqrt{2\pi\sigma^2}}$ is the right factor to ensure that the Gaussian function is a density: we compute

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2}{\sigma^2}}e^{-\frac{y^2}{\sigma^2}}dxdy = \int_{-\infty}^{\infty} e^{-\frac{x^2}{\sigma^2}}dx \cdot \int_{-\infty}^{\infty} e^{-\frac{y^2}{\sigma^2}}dy =$$

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{\sigma^2}}dx\right)^2$$

(obviously), by going to polar coordinates, and find that the double integral is equal to

$$2\pi\sigma^2$$

(just follow the previous argument). Equating we get the desired result. Incidentally, while there is no way to compute the antiderivative of $e^{-\frac{x^2}{2}}$ explicitly, in terms of "elementary" functions, the improper integral can be computed directly, without recourse to this sneaky move to two dimensions - but we need contour integration in the complex plane to do that, and that will have to wait for your complex analysis class.
After all this fluff, we go back to the actual problem:

1. If R is the distance from the center where the shot lands, $R > 4$ is the same as $R^2 > 16$. Since R^2 is assumed to be exponential with parameter **4**.

$$P\left[R^2 > 16\right] = e^{-4\cdot 16} = e^{-64} = 1.6 \cdot 10^{-28} \simeq 0$$

2. We are now looking at

$$P\left[R^2 \leq .001\right] = 1 - e^{-.01 \cdot 4} = 1 - e^{-.04} = .039211$$

or almost 4%.

3. We are wondering whether the *minimum* of two exponentials is less than .01. Since both have parameter 4, the minimum has parameter $4 + 4 = 8$ (see above, e.g., problem 4.1). Hence, the probability of hitting the center within .01 will be

$$1 - e^{-.01 \cdot 8} = 1 - e^{-.08} = .076884$$

4.6

It can be shown that if two RVs are independent and normal, their *difference* is normal with μ equal to the *difference* of their expected values, and σ^2 equal to the *sum* of their variances (you cannot reduce the variance by subtracting independent variables!). Suppose that two friends arrive at a meeting place at times, measured, in minutes, from the agreed appointment, that are independent, and distributed as normals with mean zero, and variances, respectively, 4 and 9.

1. What is the probability that either one will wait for more than 2 minutes?
2. What is the probability that they will be both on time, where "on time" is intended as within $\frac{1}{2}$ minutes from the agreed appointment?

Solution: Call the two arrival times X and Y. Note that when $X - Y < 0$ it is the second friend that waits, while if $X - Y > 0$ it is the first

1. In any case, the wait exceeds 2 minutes if $|X - Y| > 2$. Since we are dealing with a normal $N(0, 13)$, we have that

$$\frac{X - Y}{\sqrt{13}} \sim N(0, 1)$$

Our question is then

$$P\left[|Z| > \frac{2}{\sqrt{13}}\right]$$

for a standard normal Z. Since $\frac{2}{\sqrt{13}} = .55470$, this is

$$2 \cdot \left(1 - \Phi\left(\frac{2}{\sqrt{13}}\right)\right) = .57910$$

2. Now, we are looking at $P\left[|X| \leq .5, |Y| \leq .5\right] = P\left[|X| \leq .5\right] P\left[|Y| \leq .5\right]$, hence

$$P\left[\left|\frac{X}{2}\right| \leq \frac{.5}{2}\right] P\left[\left|\frac{Y}{3}\right| \leq \frac{.5}{3}\right] = 2 \cdot \left(\Phi\left(\frac{1}{4}\right) - \frac{1}{2}\right) \cdot 2 \cdot \left(\Phi\left(\frac{1}{6}\right) - \frac{1}{2}\right) =$$

$$\left(2\Phi\left(\frac{1}{4}\right)-1\right)\left(2\Phi\left(\frac{1}{6}\right)-1\right)=.026131$$

Doesn't look too great, eh?

4.7

Suppose that the distributions of trees in a forest is modeled as follows: the total area A of the forest is divided in n (n very large) small rectangles, and the probability of finding a tree in a rectangle is $\frac{\lambda}{n}$, for some constant λ (given the size of these rectangles, the probability of finding two trees in one of them is zero).

1. What is the exact distribution (according to this model) of N, the number of trees in the forest? Write the pmf.
2. If n is large enough to allow for our standard approximation, what is the probability that $N = 2\lambda$?

Solution: This is a simple Bernoulli, repeated experiment, if we assume (it should have been spelled out, because it is certainly a very unrealistic assumption!) that the events "tree in one place", and "tree in any other place" are independent. Of course, since we are not given the *joint* distribution for finding a tree in one place or another, we can do little more than assume independence, but you should realize that this is a very poor model. In fact, in real applications, we would include some "interaction", meaning that, due to pollination or other mechanisms, there is a different conditional probability of finding a tree in a spot, depending on whether there are trees nearby or not.

1. We are assuming (implicitly - sorry about that!) that the distribution of the number of trees is binomial, with parameters $n, \frac{\lambda}{n}$. Hence the pmf is

$$P[N=k]\begin{pmatrix} n \\ k \end{pmatrix}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \tag{5}$$

2. If n is large enough, we will approximate (5) with a Poisson distribution, with parameter $n \cdot \frac{\lambda}{n} = \lambda$. Hence,

$$P[N=2\lambda]=\frac{\lambda^{2\lambda}}{(2\lambda)!}e^{-\lambda}$$

(we have to hope that $2 \cdot \lambda$is an integer - if it isn't we could round to the nearest integer, or something like that).

Note It might be fun to note that, by Stirling's Formula (a version is on p. 42 of the book) $n! \simeq \sqrt{2\pi n}n^n e^{-n}$ (for n large - otherwise an additional factor of the form $e^{\frac{\theta(n)}{12n}}$ with $0 < \theta(n) < 1$ is needed. Incidentally, the "$\simeq$" has a specific meaning, but we'll just take it as "almost equal" here). Hence, in the same approximation,

$$P\left[N=2\lambda\right] \simeq \frac{\lambda^{2\lambda}}{\sqrt{2\pi\cdot 2\lambda}}\left(2\lambda\right)^{-2\lambda} e^{2\lambda}e^{-\lambda} =$$

$$= \frac{e^{\lambda}}{\sqrt{4\pi\lambda}}$$