Solution Manual for:
Applied Probability Models with Optimization Applications
by Sheldon M. Ross.

John L. Weatherwax*

November 14, 2007

Introduction

Chapter 1: Introduction to Stochastic Processes

Chapter 1: Problems

Problem 1 (the variance of $X + Y$)

We are asked to consider $\mathrm{Var}(X + Y)$ which by definition is given by

$$
\begin{aligned}
\mathrm{Var}(X+Y) &= E[((X+Y) - E[X+Y])^2] \\
&= E[(X+Y-E[X]-E[Y])^2] \\
&= E[(X-E[X]+Y-E[Y])^2] \\
&= E[(X-E[X])^2 + 2E[(X-E[X])(Y-E[Y])] + E[(Y-E[Y])^2] \\
&= \mathrm{Var}(X) + 2\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)\,,
\end{aligned}
$$

as we were asked to show.

*wax@alum.mit.edu

Problem 2 (moments of X in terms of its characteristic function)

From the definition of the characteristic function of a random variable X

$$\phi(u) = E[e^{iuX}],$$

we see that the first derivative of ϕ with respect to u is given by

$$\phi'(u) = E[iXe^{iuX}].$$

Evaluating this expression at $u = 0$ and dividing by i we find that

$$\frac{\phi'(0)}{i} = E[X].$$

Now taking the n-th derivative of $\phi(u)$ we have that

$$\frac{d^n\phi}{du^n} = E[i^n X^n e^{iuX}].$$

Further evaluating this expression at $u = 0$ we have that

$$\frac{1}{i^n}\frac{d^n\phi(0)}{du^n} = E[X^n],$$

as we were asked to show.

Problem 3 (the sum of two Poisson random variables)

We can evaluate the distribution of $X+Y$ by computing the characteristic function of $X+Y$. Since X and Y are both Poisson random variables the characteristic functions of $X+Y$ is given by

$$\begin{aligned} \phi_{X+Y}(u) &= \phi_X(u)\phi_Y(u) \\ &= e^{\lambda_1(e^{iu}-1)}e^{\lambda_2(e^{iu}-1)} \\ &= e^{(\lambda_1+\lambda_2)(e^{iu}-1)}. \end{aligned}$$

From the direct connection between characteristic functions to and probability density functions we see that the random variable $X+Y$ is a Poisson random variable with parameter $\lambda_1+\lambda_2$, the sum of the Poisson parameters of the random variables X and Y.

Problem 4 (the distribution of X given $X+Y$)

We want the conditional distribution of X given $X+Y$. Define the random variable Z by $Z = X+Y$. Then from Bayes' rule we find that

$$p(X|Z) = \frac{p(Z|X)p(X)}{p(Z)}.$$

We will evaluate each expression in tern. Now $p(X)$ is the probability density function of a Poisson random variable with parameter λ_1 so $p(X = x) = \frac{e^{-\lambda_1}\lambda_1^x}{x!}$. From problem 3 in this chapter we have that $P(Z = z) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}$. Finally to evaluate $p(Z = z|X = x)$ we recognize that this is equivalent to $p(Y = z - x)$, which we can evaluate easily. We have that

$$p(Z = z|X = x) = p(Y = z - x) = \frac{e^{-\lambda_2}\lambda_2^{z-x}}{(z-x)!} .$$

Putting all of these pieces together we find that

$$\begin{aligned} p(X = x|Z = z) &= \left(\frac{e^{-\lambda_2}\lambda_2^{z-x}}{(z-x)!}\right)\left(\frac{e^{-\lambda_1}\lambda_1^{x}}{x!}\right)\left(\frac{z!}{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}\right) \\ &= \left(\frac{z!}{x!(z-x)!}\right)\frac{\lambda_1^x\lambda_2^{z-x}}{(\lambda_1+\lambda_2)^z} \\ &= \left(\begin{array}{c} z \\ x \end{array}\right)\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{z-x} . \end{aligned}$$

Defining $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$ and $q = 1 - p = \frac{\lambda_2}{\lambda_1+\lambda_2}$ our density above becomes

$$p(X = x|Z = z) = \left(\begin{array}{c} z \\ x \end{array}\right) p^x(1-p)^{z-x} ,$$

or in words $p(X = x|Z = z)$ is a Binomial random variable with parameters $(n, p) = (z, \frac{\lambda_1}{\lambda_1+\lambda_2})$.

Problem 5 (a Poisson random variable with a random rate)

We are told that X is a Poisson random variable with parameter λ, but λ is itself an exponential random variable with mean $1/\mu$. The first statement means that

$$p(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} .$$

and $p(\lambda) = \mu e^{-\mu\lambda}$. We want to compute $p(x)$, which we can do by conditioning on all possible values of λ. This breaks the problem up into two expression that we know. We have

$$p(x) = \int p(x, \lambda)d\lambda = \int p(x|\lambda)p(\lambda)d\lambda .$$

With the definitions above we have that $p(x)$ is given by

$$\begin{aligned} p(x) &= \int_{\lambda=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}\mu e^{-\mu\lambda}d\lambda \\ &= \frac{\mu}{x!}\int_{\lambda=0}^{\infty} e^{-(\mu+1)\lambda}\lambda^x d\lambda . \end{aligned}$$

Letting $v = (1+\mu)\lambda$, so that $dv = (1+\mu)d\lambda$ the above integral becomes

$$\begin{aligned} p(x) &= \frac{\mu}{x!}\int_0^\infty e^{-v}\frac{v^x}{(1+\mu)^x}\frac{dv}{(1+\mu)} \\ &= \frac{\mu}{x!(1+\mu)^{x+1}}\int_0^\infty e^{-v}v^x dv \\ &= \frac{\mu}{x!(1+\mu)^{x+1}}x! \\ &= \frac{\mu}{(1+\mu)^{x+1}}\,. \end{aligned}$$

Defining $p = \frac{1}{1+\mu}$ and

$$q = 1 - p = 1 - \frac{1}{1+\mu} = \frac{1+\mu-1}{1+\mu} = \frac{\mu}{1+\mu}\,,$$

our density above becomes

$$p(x) = \left(\frac{1}{1+\mu}\right)^x \left(\frac{\mu}{1+\mu}\right) = p^x q\,,$$

or in words $p(x)$ is a geometric random variable with parameter $p = \frac{1}{1+\mu}$.

Problem 6 (marking chips in an urn)

Let N be the random variable denoting the draw where a previously marked (or colored) chip is drawn. We see that $P\{N=1\} = 0$, since no chip has been selected previously when the first draw is performed. Also $P\{N=2\} = 1/n$, since only one chip is marked from the n total in the urn. We can construct other probabilities by following the same logic. We find that

$$P\{N=3\} = 1\left(1-\frac{1}{n}\right)\left(\frac{2}{n}\right)\,.$$

Which can be reasoned as follows. In the first draw there are no marked chips so with probability one we will not draw a colored chip on the first draw. On the second draw only one chip is marked so we will not draw a marked chip with probability $\left(1-\frac{1}{n}\right)$. Finally, since $N=3$ we must draw a marked chip on the third draw which happens with probability $\frac{2}{n}$. For $N=4$ we find that

$$P\{N=4\} = 1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(\frac{3}{n}\right)\,.$$

which can be reasoned as before. In general we find that

$$P\{N=k\} = 1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)\cdots\left(1-\frac{k-2}{n}\right)\left(\frac{k-1}{n}\right)\,.$$

Where this expression is valid for $1 \le k \le n+1$. We can at least check that this result is a valid expression to represent a probability by selecting a value for n and verifying that when we sum the above over k for $1 \le k \le n+1$ we sum to one. A verification of this can be found in the Matlab file `chap_1_prob_6.m`.

Problem 7 (trying keys at random)

Part (a): If unsuccessful keys are removed as we try them, then the probability that the k-th attempt opens the door can be computed by recognizing that all attempts up to (but not including) the k-th have resulted in failures. Specifically, if we let N be the random variable denoting the attempt that opens the door we see that

$$
\begin{aligned}
P\{N=1\} &= \frac{1}{n} \\
P\{N=2\} &= \left(1-\frac{1}{n}\right)\frac{1}{n-1} \\
P\{N=3\} &= \left(1-\frac{1}{n}\right)\left(1-\frac{1}{n-1}\right)\frac{1}{n-2} \\
&\vdots \\
P\{N=k\} &= \left(1-\frac{1}{n}\right)\left(1-\frac{1}{n-1}\right)\left(1-\frac{1}{n-2}\right)\cdots\left(1-\frac{1}{n-(k-2)}\right)\frac{1}{n-(k-1)}\,.
\end{aligned}
$$

The expectation would be computed using the standard formula $E[N] = \sum_{k=1}^{n} kP\{N=k\}$. We can at least check that this result is a valid expression to represent a probability by selecting a value for n and verifying that when we sum the above over k for $1 \le k \le n$ we sum to one. A verification of this can be found in the Matlab file `chap_1_prob_7.m`, along with explicit calculations of the mean and variance of N.

Part (b): If unsuccessful keys are not removed then the probability that the correct key is selected at draw k is a geometric random with parameter $p = 1/n$. Thus our probabilities are given by $P\{N=k\} = (1-p)^{k-1}p$, and we have an expectation and a variance given by

$$
\begin{aligned}
E[N] &= \frac{1}{p} = n \\
\mathrm{Var}(N) &= \frac{1-p}{p^2} = n(n-1)\,.
\end{aligned}
$$

Problem 8 (the variance of Y in terms of the conditional variance)

We begin by recalling the definition of the conditional variance which is given by

$$\mathrm{Var}(Y|X) = E[(Y - E[Y|X])^2|X]\,.$$

In this problem we want to show that

$$\mathrm{Var}(Y) = E[\mathrm{Var}(Y|X)] + \mathrm{Var}(E[Y|X])\,.$$

Consider the definition of $\mathrm{Var}(Y|X)$ by expanding the quadratic. We find that

$$
\begin{aligned}
\mathrm{Var}(Y|X) &= E[Y^2 - 2YE[Y|X] + E[Y|X]^2|X] \\
&= E[Y^2|X] - 2E[YE[Y|X]|X] + E[E[Y|X]^2|X]\,.
\end{aligned}
$$

Now $E[Y|X]$ is a function of X (since it is defined as $\int yp(y|x)dy$) so the nested expectation in the second term above becomes

$$E[YE[Y|X]|X] = E[Y|X]E[Y|X] = E[Y|X]^2\,,$$

while the third term in the above becomes

$$E[E[Y|X]^2|X] = E[Y|X]^2 E[1|X] = E[Y|X]^2\,.$$

So will all of these substitutions the expression for $\mathrm{Var}(Y|X)$ above becomes

$$\begin{aligned}\mathrm{Var}(Y|X) &= E[Y^2|X] - 2E[Y|X]^2 + E[Y|X]^2 \\ &= E[Y^2|X] - 2E[Y|X]^2 + E[Y|X]^2 \\ &= E[Y^2|X] - E[Y|X]^2\,.\end{aligned}$$

Taking the expectation of the above with respect to X we have

$$\begin{aligned}E[\mathrm{Var}(Y|X)] &= E[E[Y^2|X]] - E[E[Y|X]^2] \\ &= E[Y^2] - E[E[Y|X]^2]\,,\end{aligned}$$

using the properties of nested expectation. Subtracting $E[Y]^2$ from both sides and moving $E[E[Y|X]^2]$ to the left hand side we obtain

$$E[Y^2] - E[Y]^2 = E[\mathrm{Var}(Y|X)] + E[E[Y|X]^2] - E[Y]^2\,.$$

Now the left hand side of the above is equivalent to $\mathrm{Var}(Y)$ and $E[Y]^2 = E[E[Y|X]]^2$ (using the properties of nested expectation). With these two substitutions the expression for $\mathrm{Var}(Y)$ above becomes

$$\begin{aligned}\mathrm{Var}(Y) &= E[\mathrm{Var}(Y|X)] + E[E[Y|X]^2] - E[E[Y|X]]^2 \\ &= E[\mathrm{Var}(Y|X)] + \mathrm{Var}(E[Y|X])\,.\end{aligned}$$

Problem 9 (the expected amount spent)

We will assume that for this problem that N (the number of customers) is a constant and not a random variable. Let X_i be the random variable denoting the amount spent by the ith customer. Then the total amount of money spent in the store is given by

$$T = \sum_{i=1}^{N} X_i\,.$$

So the expected amount of money is given by

$$E[T] = \sum_{i=1}^{N} E[X_i] = E[X_i]\cdot N$$

Where $E[X_i] = \int x dF(x)$, and $F(x)$ is the cumulative distribution function for the random variables X_i. Since we assume that the variables X_i are independent we have

$$\mathrm{Var}(T) = \sum_{i=1}^{N} \mathrm{Var}(X_i) = N\mathrm{Var}(X_i)\,.$$

Here $\mathrm{Var}(X_i)$ is given by the usual expression

$$\mathrm{Var}(X_i) = \int x^2 dF(x) - \left(\int x dF(x)\right)^2\,.$$

Problem 10 (an independent increments process is also a Markov process)

Let assume that our stochastic process $\{X(t), t \in T\}$ has independent increments. This means that for all choices of times $t_0 < t_1 < t_2 < \cdots < t_n$ then the n random variables

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})\,,$$

are independent. We desire to show that $\{X(t), t \in T\}$ is then a Markov process. The definition of a Markov process is that for all sequences of times $t_0 < t_1 < \cdots < t_{n-1} < t$ and process values $x_0 < x_1 < \cdots < x_{n-1} < x$ we have the following identity on the cumulative probability density of $X(t)$ conditioned on the total sample path

$$P\{X(t) \le x | X(t_0) = x_0, X(t_1) = x_1, \cdots, X(t_{n-1}) = x_{n-1}\} = P\{X(t) \le x | X(t_{n-1}) = x_{n-1}\}\,.$$

To show this, for our independent increments process consider the expression

$$P\{X(t) \le x | X(t_0) = x_0, X(t_1) = x_1, \cdots, X(t_{n-1}) = x_{n-1}\}\,,$$

which we can write (introducing increment variables) as follows

$$\begin{array}{rll} P\{X(t) - X(t_{n-1}) \le x - x_{n-1} | & X(t_{n-1}) & -X(t_{n-2}) = x_{n-1} - x_{n-2}, \\ & X(t_{n-2}) & -X(t_{n-3}) = x_{n-2} - x_{n-3}, \\ & \cdots & \\ & X(t_2) & -X(t_1) = x_2 - x_1, \\ & X(t_1) & -X(t_0) = x_1 - x_0, \\ & X(t_0) & -X(0) = x_0 - X(0)\} \end{array}$$

For any $X(0) < x_0$. Then the property of independent increments states that this expression is equal to

$$P\{X(t) - X(t_{n-1}) \le x - x_{n-1}\}\,.$$

This is because each random variable (in the conditional expression) is an increment random variable and our stochastic process is an independent increments process. To finish this problem simply requires that we next recognize that the above probability is equivalent to

$$P\{X(t) \le x - x_{n-1} + X(t_{n-1})\} = P\{X(t) \le x | X(t_{n-1}) = x_{n-1}\}\,.$$

Since this last expression is equivalent to the definition of a Markov process we have shown that our independent increments process is a Markov process.

Problem 11 (the probability all samples are greater than x)

Part (a): Now $P\{Z_n > x\} = 1 - P\{Z_n \le x\}$. Since $P\{Z_n \le x\}$ is the cumulative distribution function for Z_n (defined as $Z_n = \min(Y_{1,n}, Y_{2,n}, \cdots, Y_{n,n})$) we can evaluate the cumulative distribution function by first noting that the distribution function for Z_n is given by

$$f_{Z_n}(z) = n(1 - F(z))^{n-1} f(z) .$$

Where $f(\cdot)$, and $F(\cdot)$ is the distribution and cumulative distribution function for the random variables $Y_{n,i}$. For a derivation of this result see [1]. In our case $Y_{n,i}$ is a uniform random variable with range $(0, t)$ and so $f(z) = \frac{1}{t}$, and $F(z) = \frac{z}{t}$. We thus have that

$$f_{Z_n}(z) = n\left(1 - \frac{z}{t}\right)^{n-1} \frac{1}{t} ,$$

so that the cumulative distribution function for Z_n is given by

$$\begin{aligned} P\{Z_n \le x\} &\equiv \int_0^x f_{Z_n}(\xi) d\xi = \int_0^x n\left(1 - \frac{\xi}{t}\right)^{n-1} \frac{1}{t} \\ &= 1 - \left(1 - \frac{x}{t}\right)^n . \end{aligned}$$

So that

$$P\{Z_n > x\} = 1 - \left(1 - \left(1 - \frac{x}{t}\right)^n\right) = \left(1 - \frac{x}{t}\right)^n .$$

Part (b): Assuming that t is a function of n ($t = t(n)$) such that

$$\lim_{n\to\infty} \frac{n}{t(n)} = \lambda ,$$

we can evaluate the requested limit as follows

$$\begin{aligned} \lim_{n\to\infty} P\{Z_n > x\} &= \lim_{n\to\infty} \left(1 - \frac{x}{t}\right)^n \\ &= \lim_{n\to\infty} \left(1 - \frac{\left(\frac{n}{t}\right) x}{n}\right)^n . \end{aligned}$$

To further evaluate this expression, we remember the following famous limit from calculus

$$\lim_{n\to\infty} \left(1 + \frac{\xi}{n}\right)^n = e^{\xi} ,$$

from which we see that the above limit is equal to $e^{-\lambda x}$ as claimed.

Problem 12 (continuous solutions to two functional equations)

Part (a): Consider the functional equation 6 which is $f(t + s) = f(t) f(s)$. We will solve this problem by computing $f(\cdot)$ at an increasing number of points on the real line and then finally concluding what the function $f(\cdot)$ must be. We begin by computing f at zero. Letting $t = s = 0$ in our functional equation we find that

$$f(0) = f(0)^2 \quad \text{or} \quad f(0) = 1\,.$$

We will now show that to compute $f(t)$ for t an positive integer. We begin by letting $s = 1$, we find that $f(t+1) = f(1)f(t)$, in the same way we have that $f(t+2)$ is given by $f(t+2) = f(1)f(t+1) = f(1)^2 f(t)$. Generalizing this process we see that for any positive integer s we have that

$$f(t+s) = f(1)^s f(t)\,.$$

At the same time these manipulations can be used to compute f for negative integers. We find that

$$f(t-s) = f(1)^{-s} f(t)\,.$$

From which if $t = 0$ we find that $f(s) = f(1)^s$, and $f(-s) = f(1)^{-s}$. In both of these expressions the value of $f(1)$ is an unknown constant. The benefit of deriving these expressions is that we see that to evaluate f in any interval of the real line can be reduced to an equivalent problem of evaluating $f(\cdot)$ on the interval $0 \le t \le 1$, and multiplying this result by a power of $f(1)$. Thus we now attempt to evaluate $f(\cdot)$ for rational numbers s and t such that $0 \le s, t \le 1$. To begin lets consider $s = \frac{1}{2}$ and $t = \frac{1}{2}$. We find that

$$f(1) = f(\frac{1}{2})f(\frac{1}{2}) \quad \text{so} \quad f(\frac{1}{2}) = f(1)^{\frac{1}{2}}\,.$$

Letting $t = \frac{1}{3}$ and $s = \frac{2}{3}$ we find that

$$f(1) = f(\frac{1}{3})f(\frac{2}{3}) = f(\frac{1}{3})f(\frac{1}{3}+\frac{1}{3}) = f(\frac{1}{3})f(\frac{1}{3})f(\frac{1}{3}) = f(\frac{1}{3})^3 \quad \text{so} \quad f(\frac{1}{3}) = f(1)^{\frac{1}{3}}\,.$$

Continuing this train of thought we see that in general then

$$f(\frac{1}{n}) = f(1)^{\frac{1}{n}} \quad \text{for} \quad n > 0\,,$$

and thus we have evaluated $f(\cdot)$ at these specific fractions. Using our functional relationship we can evaluate f at the fractions $\frac{k}{n}$ as follows

$$f(\frac{k}{n}) = f(\frac{1}{n})f(\frac{k-1}{n}) = f(\frac{1}{n})f(\frac{1}{n})\cdots f(\frac{1}{n}) = f(\frac{1}{n})^k = f(1)^{\frac{k}{n}}\,.$$

And thus we have evaluated f for a rational values $r = \frac{k}{n}$. Using the known continuity of f we can then conclude that for any real x,

$$f(x) = f(1)^x\,.$$

Now defining λ such that $f(1) = e^{-\lambda}$ which is equivalent to $\lambda = -\ln(f(1))$ our function f becomes (in terms of λ)

$$f(x) = e^{-\lambda x}\,,$$

as expected.

Part (b): Consider the functional equation 7 which is $f(t+s) = f(t) + f(s)$. We will solve this problem by computing $f(\cdot)$ at an increasing number of points on the real line and then finally concluding what the function $f(\cdot)$ must be. We begin by computing f at zero. Letting $t = s = 0$ in our functional equation we find that

$$f(0) = f(0) + f(0) \quad \text{or} \quad f(0) = 0\,.$$

We will now show that to compute $f(t)$ for t an positive integer. Consider

$$f(t) = f(t-1+1) = f(t-1) + f(1) = f(t-2) + 2f(1) = \cdots = tf(1)\,.$$

We can compute f for negative t by letting t be positive and $s = -t$ so that our functional equation 7 becomes

$$f(0) = f(t) + f(-t)\,,$$

or since $f(0) = 0$ we then have that

$$f(-t) = -f(t)\,.$$

We thus have expressions for f for all integer x, in terms of an unknown constant $f(1)$. As in Part (a) of this problem if we desire to evaluate f at an x that is not in the interval $0 < x < 1$, using the functional relationships possessed by f we can simplify the problem to one where we only need to evaluate f inside the interval $(0, 1)$. To show an example of this let x be a positive real number with "integer" part n such that $x = n + \xi$, where $0 < \xi < 1$. We find that

$$\begin{aligned}
f(x) &= f(n+\xi) \\
&= f(1 + (n-1) + \xi) \\
&= f(1) + f(n-1+\xi) \\
&= 2f(1) + f(n-2+\xi) \\
&= \cdots \\
&= nf(1) + f(\xi)\,.
\end{aligned}$$

Thus we have reduced our problem to one where we only have to evaluate $f(x)$ for $0 < x < 1$. Lets begin to evaluate f in this range by expression $f(x)$ when x is rational. To begin let $t = \frac{1}{n}$, and $s = \frac{1}{n}$, for which we find that

$$f(\frac{1}{n} + \frac{1}{n}) = f(\frac{1}{n}) + f(\frac{1}{n}) = 2f(\frac{1}{n})\,.$$

In the same way we find that $f(\frac{k}{n}) = kf(\frac{1}{n})$. Letting $k = n$ in this expression gives $f(\frac{1}{n}) = \frac{f(1)}{n}$ so that

$$f(\frac{k}{n}) = kf(\frac{1}{n}) = \frac{k}{n}f(1)\,.$$

This result shows that for x rational we have $f(x) = xf(1)$. Using the continuity of f we have that for *real* x, that $f(x) = xf(1)$. Defining $c = f(1)$ we see that this is equivalent to $f(x) = cx$, as we were asked to show.

Problem 13 (the variance of the Wiener process)

A Wiener process is defined as a stochastic process $\{X(t), t \geq 0\}$ if:

i $\{X(t), t \geq 0\}$ has stationary, independent increments.

ii For every $t \geq 0$, $X(t)$ is normally distributed with mean 0.

iii $X(0) = 0$.

Now introduce the function $f(t)$ as $f(t) \equiv \text{Var}(X(t))$, and consider the expression $f(t+s)$. We find that

$$\begin{aligned} f(t+s) &= \text{Var}(X(t+s)) \\ &= \text{Var}((X(t+s) - X(s)) + (X(s) - X(0)))\,. \end{aligned}$$

Where we have explicity introduced increment variables so that we can take advantage of the increments properties of the Wiener process. Specifically since $X(\cdot)$ has independent increments the above can be written as

$$\text{Var}(X(t+s) - X(s)) + \text{Var}(X(s) - X(0))\,.$$

Now because $X(\cdot)$ is stationary, we know that the distribution of the random variables $X(t_2 + s) - X(t_1 + s)$ and $X(t_2) - X(t_1)$ is the same, the variance

$$\text{Var}(X(t+s) - X(s)) = \text{Var}(X(t) - X(0)) = \text{Var}(X(t)) = f(t)\,.$$

So that the function $f(\cdot)$ must satisfy the following functional equation

$$f(t+s) = f(t) + f(s)\,.$$

From Problem 12 in this chapter the unique continuous solution to this function equation is $f(t) = ct$. Since $\text{Var}(X(t)) > 0$, we can assume our constant c is positive, and take it to be σ^2, so that $f(t) = \sigma^2 t$, as expected.

Chapter 2: The Poisson Process

Chapter 2: Problems

Problem 1 (equivalent definitions of a Poisson process)

We are asked to prove the equivalence of two definitions for a Poisson process. The first definition (Definition 2.1 in the book) is the following

The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process if:

i $N(0) = 0$

ii $\{N(t), t \geq 0\}$ has independent increments

iii The number of events in any interval of length t has a Poisson distribution with mean λt. That is for $s, t \geq 0$, we have

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad n \geq 0$$

We want to show that this definition is equivalent to the following (which is Definition 2.2 in the book)

i $N(0) = 0$

ii $\{N(t), t \geq 0\}$ has stationary, independent increments

iii $P\{N(t) \geq 2\} = o(t)$

iv $P\{N(t) = 1\} = \lambda t + o(t)$

We begin by noting that both definitions require $N(0) = 0$. From (ii) in Definition 2.1 we have the required independent increments needed in Definition 2.2 (ii). From (iii) in Definition 2.1 we have that the distributions of $X(t_2 + s) - X(t_1 + s)$ is given by a Poisson distribution with mean $\lambda(t_2 - t_1)$ and the distribution of random variable $X(t_2) - X(t_1)$ is also given by a Poisson distribution with mean $\lambda(t_2 - t_1)$ showing that the process $\{N(t)\}$ also has *stationary* increments and thus satisfies the totality of Definition 2.2 (ii).

From (iii) in Definition 2.1 we have with $s = 0$ (and the fact that $N(0) = 0$) that

$$P\{N(t) = n\} = \frac{e^{-\lambda t}(\lambda t)^n}{n!}.$$

So that

$$\begin{aligned}
P\{N(t) \geq 2\} &= \sum_{n=2}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\left[\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} - 1 - \lambda t\right] \\
&= e^{-\lambda t}\left[e^{\lambda t} - 1 - \lambda t\right] \\
&= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t},
\end{aligned}$$

which (we claim) is a function that is $o(t)$. To show that this is true consider the limit as t goes to zero. Thus we want to evaluate

$$\lim_{t \to 0} \frac{1 - e^{-\lambda t} - \lambda t e^{-\lambda t}}{t}.$$

Since this is an indeterminate limit of type $0/0$ we must use L'Hopital's rule which gives that the above limit is equal to the limit of the derivative of the top and bottom of the above or

$$\lim_{t\to 0}\frac{\lambda e^{-\lambda t}-\lambda e^{-\lambda t}+\lambda^2 te^{-\lambda t}}{1}=\lambda-\lambda=0\,.$$

Proving that this expression is $o(t)$ (since this limit equaling zero is the definition) and proving that $P\{N(t)\geq 2\}=o(t)$. The final condition required for Definition 2.2 is (iv). We have from Definition 2.1 (iii) that

$$P\{N(t)=1\}=\frac{e^{-\lambda t}(\lambda t)}{1!}=\lambda te^{-\lambda t}$$

To show that this expression has the correct limiting behavior as $t\to 0$, we first prove that

$$e^{-\lambda t}=1-\lambda t+o(t)\quad\text{as}\quad t\to 0\,,$$

Which we do by evaluating the limit

$$\lim_{t\to 0}\frac{e^{-\lambda t}-1+\lambda t}{t}=\lim_{t\to 0}\frac{-\lambda e^{-\lambda t}+\lambda}{1}=-\lambda+\lambda=0\,.$$

Where we have used L'Hopital's rule again. With this result we see that

$$\begin{aligned}P\{N(t)=1\} &= \lambda t(1-\lambda t+o(t))\\ &= \lambda t-\lambda^2t^2+o(t^2)\\ &= \lambda t+o(t)\,,\end{aligned}$$

showing the truth of condition (iv) in Definition 2.2.

Problem 2 (we can derive the fact that $P\{N(t)=1\}=\lambda t+o(t)$)

Following the hint for this problem we will try to derive a functional relationship for $P\{N(t)=0\}$, by considering $P\{N(t+s)=0\}$. Now if $N(t+s)=0$, this event is equivalent to the event that $N(t)=0$ *and* $N(t+s)-N(t)=0$. so we have that

$$\begin{aligned}P\{N(t+s)=0\} &= P\{N(t)=0,N(t+s)-N(t)=0\}\\ &= P\{N(t)-N(0)=0,N(t+s)-N(t)=0\}\\ &= P\{N(t)-N(0)=0\}P\{N(t+s)-N(t)=0\}\\ &= P\{N(t)=0\}P\{N(s)=0\}\,.\end{aligned}$$

When we used the property of stationary independent increments. Thus defining $f(t)\equiv P\{N(t)=0\}$, from the above we see that f satisfies

$$f(t+s)=f(t)f(s)\,.$$

By the discussion in the book the unique continuous solution to this equation is $f(t)=e^{-\lambda t}$, for some λ. Thus we have that $P\{N(t)=0\}=e^{-\lambda t}$. Using (iii) from Definition 2.2 and the fact that probabilities must be normalized (sum to one) we have that

$$P\{N(t)=0\}+P\{N(t)=1\}+P\{N(t)\geq 2\}=1\,.$$

which gives us (solving for $P\{N(t)=1\}$) the following

$$
\begin{aligned}
P\{N(t)=1\} &= 1-P\{N(t)=0\}-P\{N(t)\geq 2\} \\
&= 1-e^{-\lambda t}-o(t) \\
&= 1-(1-\lambda t+o(t))-o(t) \\
&= \lambda t+o(t)\,,
\end{aligned}
$$

as we were to show.

Problem 3 (events registered with probability p)

Let $M(t)$ be the process where we register each event i from a Poisson process (with rate λ) with probability p. Then we want to show that $M(t)$ is another Poisson process with rate $p\lambda$. To do so consider the probability that $M(t)$ has counted j "events", by conditioning on the number of observed events from the original Poisson process. We find

$$
P\{M(t)=j\}=\sum_{n=0}^{\infty}P\{M(t)=j|N(t)=n\}\frac{e^{-\lambda t}(\lambda t)^n}{n!}
$$

The conditional probability in this sum can be computed using the rule defined above since if we have n original events the number of derived events is a binomial random variable with parmeters (n,p). Specifically then we have

$$
P\{M(t)=j|N(t)=n\}=\begin{cases}\left(\begin{array}{c} n \\ j \end{array}\right)p^j(1-p)^{n-j} & j\leq n \\ 0 & j>n\,.\end{cases}
$$

Putting this result into the original expression for $P\{M(t)=j\}$ we find that

$$
P\{M(t)=j\}=\sum_{n=j}^{\infty}\left(\begin{array}{c} n \\ j \end{array}\right)p^j(1-p)^{n-j}\left(\frac{e^{-\lambda t}(\lambda t)^n}{n!}\right)
$$

To evaluate this we note that $\left(\begin{array}{c} n \\ j \end{array}\right)\frac{1}{n!}=\frac{1}{j!(n-j)!}$, so that the above simplifies as following

$$
\begin{aligned}
P\{M(t)=j\} &= \frac{e^{-\lambda t}p^j}{j!}\sum_{n=j}^{\infty}\frac{1}{(n-j)!}(1-p)^{n-j}(\lambda t)^n \\
&= \frac{e^{-\lambda t}p^j}{j!}\sum_{n=j}^{\infty}\frac{1}{(n-j)!}(1-p)^{n-j}(\lambda t)^j(\lambda t)^{n-j} \\
&= \frac{e^{-\lambda t}(p\lambda t)^j}{j!}\sum_{n=j}^{\infty}\frac{((1-p)\lambda t)^{n-j}}{(n-j)!} \\
&= \frac{e^{-\lambda t}(p\lambda t)^j}{j!}\sum_{n=0}^{\infty}\frac{((1-p)\lambda t)^{n}}{n!} \\
&= \frac{e^{-\lambda t}(p\lambda t)^j}{j!}e^{(1-p)\lambda t}=e^{-p\lambda t}\frac{(p\lambda t)^j}{j!}\,,
\end{aligned}
$$

from which we can see $M(t)$ is a Poisson process with rate λp.

Problem 4 (the correlation of a Poisson process)

Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ. Then manipulating the expression we desire into increment variables and remembering that $N(0) = 0$, we find

$$\begin{aligned}
E[N(t)N(t+s)] &= E[N(t)(N(t+s) - N(s) + N(s))] \\
&= E[N(t)(N(t+s) - N(s))] + E[N(t)N(s)] \\
&= E[(N(t) - N(0))(N(t+s) - N(s))] \\
&+ E[(N(t) - N(0))(N(s) - N(0))] \\
&= E[N(t) - N(0)]\, E[N(t+s) - N(s)] \\
&+ E[N(t) - N(0)]\, E[N(s) - N(0)]\,.
\end{aligned}$$

Where we have used the independent increments property of the Poisson process. Now from the fact that a Poisson process is also a stationary process

$$E[N(t+s) - N(s)] = E[N(t) - N(0)] = \lambda t\,.$$

Thus the above expression becomes

$$E[N(t)E(t+s)] = \lambda t \cdot \lambda t + \lambda t \cdot \lambda s = \lambda^2 t(t-s)\,.$$

Problem 5 (the sum of two Poisson processes)

Since $N_1(t)$ and $N_2(t)$ are both Poisson random variables with parameters $\lambda_1 t$ and $\lambda_2 t$ respectively, from Problem 3 in Chapter 1 of this book the random variable $M(t)$ defined by $N_1(t) + N_2(t)$ is a Poisson random variable with parameter $\lambda_1 t + \lambda_2 t$ and thus has a probability of the event $M(t) = j$ given by

$$P\{M(t) = j\} = \frac{e^{-(\lambda_1 t + \lambda_2 t)}(\lambda_1 t + \lambda_2 t)^j}{j!} = \frac{e^{-(\lambda_1+\lambda_2)t}((\lambda_1+\lambda_2)t)^j}{j!}\,,$$

showing that $M(t)$ is a Poisson process with rate $\lambda_1 + \lambda_2$.

For the second part of this problem we want to evaluate

$$P\{N_1(t) = 1, N_2(t) = 0 | N_1(t) + N_2(t) = 1\}\,,$$

which we can do by using the definition of conditional probabilities as

$$\begin{aligned}
P\{N_1(t) = 1, N_2(t) = 0 | N_1(t) + N_2(t) = 1\} &= \frac{P\{N_1(t) = 1, N_2(t) = 0\}}{P\{N_1(t) + N_2(t) = 1\}} \\
&= \frac{P\{N_1(t) = 1\} P\{N_2(t) = 0\}}{P\{N_1(t) + N_2(t) = 1\}}\,.
\end{aligned}$$

In the above we have used the independence of the process $N_1(\cdot)$ and $N_2(\cdot)$. The above then equals

$$\frac{\frac{e^{-\lambda_1 t}(\lambda_1 t)^1}{1!} \cdot e^{-\lambda_2 t}}{\frac{e^{-(\lambda_1+\lambda_2)t}((\lambda_1+\lambda_2)t)^1}{1!}} = \frac{\lambda_1}{\lambda_1 + \lambda_2}\,,$$

as we were to show.

Problem 6 (the probability that $N_1(t)$ reaches n before $N_2(t)$ reaches m)

Warning: I would appreciate feedback if this solution is not correct in anyway or if you agree with the solution method.

We are asked to compute the probabilty of the joint event $N_1(t) = n$ and $N_2(t) < m$, i.e.

$$P\{N_1(t) = n, N_2(t) < m\},$$

which we will evaluate by conditioning on the event $N_1(t) = n$. We have the above equal to

$$P\{N_2(t) < m | N_1(t) = n\} P\{N_1(t) = n\}.$$

Now since N_1 and N_2 are assumed independent, we have that the above equals

$$P\{N_2(t) < m\} P\{N_1(t) = n\},$$

from which each term can be evaluated. Remembering that for a Poisson procsses with rate λ, we have that

$$P\{N(t) = n\} = \frac{e^{-\lambda t}(\lambda t)^n}{n!},$$

the above is then equal to

$$P\{N_1(t) = n, N_2(t) < m\} = \left(\sum_{j=0}^{m-1} \frac{e^{-\lambda_2 t}(\lambda_2 t)^j}{j!} \right) \left(\frac{e^{-\lambda_1 t}(\lambda_1 t)^n}{n!} \right).$$

Problem 7 (the expectation and the variance of $N(Y)$)

We will compute both of these expressions by conditioning on the value of the random variable Y e.g. for the first expression $E[N(Y)] = E[\,E[N(Y)|Y]\,]$. Now since $E[N(Y)|Y = y] = \lambda y$, we find that

$$E[N(Y)] = E[\lambda Y] = \lambda E[Y].$$

For the variance calculation we have that $\text{Var}(N(Y)) = E[N(Y)^2] - E[N(Y)]^2$, from which we can evaluate the first expression by conditional expectations as earlier. We have that (remembering the definition of the variance for Poisson random variable) that

$$E[N(Y)^2 | Y = y] = \lambda y + \lambda^2 y^2,$$

so that we have upon taking the expectation with respect to Y that

$$E[N(Y)^2] = \lambda E[Y] + \lambda^2 E[Y^2].$$

This gives for the expression $\text{Var}(N(Y))$ the following

$$\begin{aligned} \text{Var}(N(Y)) &= \lambda E[Y] + \lambda^2 E[Y^2] - \lambda^2 E[Y]^2 \\ &= \lambda E[Y] + \lambda^2 \text{Var}(Y). \end{aligned}$$

Problem 8 (the infinite server Poisson queue)

From the example in the book, the probability that a customer who arrives at time x will not be present (i.e. will have completed service) at time t is given by $G(t-x)$. In the book's example the corresponding expression was $1 - G(t-x)$. Recognizing this of the manipulations carry through from that example. Thus defining

$$q \equiv \int_0^t G(t-x)\frac{dx}{t} = \int_0^t G(x)\frac{dx}{t},$$

and heavily using the example result we see that

$$P\{Y(t) = j\} = \frac{e^{-\lambda t q}(\lambda t q)^j}{j!}.$$

Thus $Y(t)$ has a Poisson distribution with a parameter $\lambda q = \lambda \int_0^t G(x)\frac{dx}{t}$. In addition, to this result we recognize that the number of customers who have *completed* service by time t is the "complement" of those that are still in the system at time t (this later number is $X(t)$). This means that our random variable $Y(t)$ is equivalent to (using the notation from that section)

$$Y(t) = N(t) - X(t).$$

With these observations we will now prove that $X(t)$ and $Y(t)$ are independent. From the discussion above we have shown that

$$P\{X(t) = i\} = \frac{e^{-\lambda t p}(\lambda t p)^i}{i!} \quad \text{and} \quad P\{Y(t) = j\} = \frac{e^{-\lambda t p}(\lambda t p)^j}{j!}.$$

With $N(t) = X(t) + Y(t)$. To investigate the independence of $X(t)$ and $Y(t)$ consider the discrete joint density over the pair of variables $(X(t), Y(t))$ i.e.

$$P\{X(t) = i, Y(t) = j\}.$$

Now using an equivalent expression for $Y(t)$ the above equals

$$P\{X(t) = i, N(t) - X(t) = j\},$$

or

$$P\{X(t) = i, N(t) = i + j\}.$$

By the definition of conditional probability we have the above equal to

$$P\{X(t) = i, N(t) = i + j\} = P\{X(t) = i | N(t) = i + j\} P\{N(t) = i + j\}.$$

We can directly compute each of these two terms. The first term $P\{X(t) = i | N(t) = i + j\}$ is the probability of an binomial random variable with probability of success p, while the second term is obtained from the fact that $N(t)$ is a Poisson process with rate λ. Thus we have the above expression equal to (using $p+q=1$)

$$\begin{aligned}
P\{X(t)=i|N(t)=i+j\}P\{N(t)=i+j\} &= \left(\begin{array}{c} i+j \\ i \end{array}\right) p^i(1-p)^j \left(\frac{e^{-\lambda t}(\lambda t)^{i+j}}{(i+j)!}\right) \\
&= \frac{(i+j)!}{i!j!} p^i(1-p)^j \left(\frac{e^{-\lambda t}(\lambda t)^{i+j}}{(i+j)!}\right) \\
&= \frac{p^i q^j}{i!j!} e^{-\lambda(p+q)t}(\lambda t)^i(\lambda t)^j \\
&= \left(\frac{(\lambda p t)^i e^{-\lambda p t}}{i!}\right)\left(\frac{(\lambda q t)^j e^{-\lambda q t}}{j!}\right) \\
&= P\{X(t)=i\}P\{Y(t)=j\}\,.
\end{aligned}$$

Thus we have shown that the joint density equals the product of the marginal densities the random variables $X(t)$ and $Y(t)$ are independent.

Problem 9 (sums of jointly varying random variables)

We are told that $S_1, S_2, \cdots$ are the arrival times for a Poisson process, i.e. as an unordered sequence of numbers these are independent identically distributed random variables with a uniform distribution function over $(0,t)$. Define the random variable $X(t)$ as

$$X(t) = \sum_{i=1}^{N(t)} g(Y_i, S_i)\,.$$

Lets begin by finding the characteristic function of this random variable. Closely following the electronic counter example from the book we see that

$$\begin{aligned}
\phi_{X(t)}(u) &= E[e^{iuX(t)}] \\
&= \sum_{n=0}^{\infty} E[e^{iuX(t)}|N(t)=n]\frac{e^{-\lambda t}(\lambda t)^n}{n!}
\end{aligned}$$

Now this internal expectation in terms of the sum of the g functions above given by

$$E[e^{iuX(t)}|N(t)=n] = E\left[\exp\left\{iu\sum_{i=1}^{n} g(Y_i,S_i)\right\}\right]\,.$$

Where this expectation is taken with respect to the variables Y_i and S_i. We know that the unordered sequence of arrival times S_i are uniformly distributed over the interval $(0,t)$. We will also assume that the cumulative distribution function for Y_i is given by $G(y)$. Then since everything is independent we have that

$$E\left[\exp\left\{iu\sum_{i=1}^{n} g(Y_i,S_i)\right\}\right] = E\left[\exp\left\{iug(Y_1,S_1)\right\}\right]^n\,.$$

This final expectation can be taken just with respect to one Y and one S variable. By definition it is equal to (assuming the domain of the y values are given by Ω_y)

$$E\left[\exp\left\{iug(Y_1, S_1)\right\}\right] = \int_{\Omega_y}\int_0^t e^{iug(y,s)}\frac{ds}{t}\,dG(y) \equiv I\,.$$

We define this expression as I to save notation. This expression can now put back into the expression for the characteristic function of X to give

$$\begin{aligned}
\phi_{X(t)}(u) &= E[e^{iuX(t)}] \\
&= \sum_{n=0}^{\infty} I^n \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(I\lambda t)^n}{n!} \\
&= e^{-\lambda t}e^{\lambda t I} = e^{-\lambda t(1-I)}\,.
\end{aligned}$$

We can evaluate $1 - I$ by recognizing that since to be properly normalized probability distributions we must have

$$1 = \int_{\Omega_y}\int_0^t \frac{ds}{t}\,dG(y)\,.$$

So that

$$1 - I = \int_{\Omega_y}\int_0^t (1 - e^{iug(y,s)})\frac{ds}{t}\,dG(y)\,.$$

When we put back in the expression for I we find the characteristic function for $X(t)$ to be given by

$$\phi_{X(t)}(u) = \exp\left\{-\lambda t\left(\int_{\Omega_y}\int_0^t (1 - e^{iug(y,s)})\frac{ds}{t}\,dG(y)\right)\right\}.$$

To evaluate the expectations we now compute $E[X(t)] = \phi'_X(0)/i$. We find that the first derivative of $\phi_X(u)$ is given by

$$\begin{aligned}
\phi'_X(u) &= \lambda t\exp\left\{-\lambda t\left(\int_{\Omega_y}\int_0^t (1 - e^{iug(y,s)})\frac{ds}{t}\,dG(y)\right)\right\} \\
&\times \left(\int_{\Omega_y}\int_0^t (ig(y,s)e^{iug(y,s)})\frac{ds}{t}\,dG(y)\right).
\end{aligned}$$

Evaluating the above for $u = 0$ and dividing by i we obtain

$$\begin{aligned}
E[X(t)] &= \lambda t\left(\int_{\Omega_y}\int_0^t g(y,s)\frac{ds}{t}\,dG(y)\right) \\
&= \lambda t E[g(Y,S)]
\end{aligned}$$

To compute the variance we need the second derivative of our characteristic function. Specifically, $\mathrm{Var}(X(t)) = -\phi_X''(0) - E[X(t)]^2$. We find this second derivative given by

$$
\begin{aligned}
\phi_X''(u) &= \lambda^2 t^2 \exp\left\{ -\lambda t \left(\int_{\Omega_y} \int_0^t (1 - e^{iug(y,s)}) \frac{ds}{t}\, dG(y) \right) \right\} \\
&\times \left(\int_{\Omega_y} \int_0^t (ig(y,s) e^{iug(y,s)}) \frac{ds}{t}\, dG(y) \right)^2 \\
&+ \lambda t \exp\left\{ -\lambda t \left(\int_{\Omega_y} \int_0^t (1 - e^{iug(y,s)}) \frac{ds}{t}\, dG(y) \right) \right\} \\
&\times \left(\int_{\Omega_y} \int_0^t (i^2 g(y,s)^2 e^{iug(y,s)}) \frac{ds}{t}\, dG(y) \right) .
\end{aligned}
$$

Which could be simplified further but with more algebra. When we evaluate this for $u = 0$ we obtain the following

$$
\begin{aligned}
\phi_X''(0) &= -\lambda^2 t^2 \left(\int_{\Omega_y} \int_0^t g(y,s) \frac{ds}{t}\, dG(y) \right)^2 - \lambda t \left(\int_{\Omega_y} \int_0^t g(y,s)^2 \frac{ds}{t}\, dG(y) \right) \\
&= -\lambda^2 t^2 E[g(Y,S)]^2 - \lambda t E[g(Y,S)^2] .
\end{aligned}
$$

With these results then $\mathrm{Var}(X(t))$ is given by

$$
\begin{aligned}
\mathrm{Var}(X(t)) &= \lambda^2 t^2 E[g(Y,S)]^2 + \lambda t E[g(Y,S)^2] - \lambda^2 t^2 E[g(Y,S)]^2 \\
&= \lambda t E[g(Y,S)^2] .
\end{aligned}
$$

Problem 11 (the expectation and variance of a compound Poisson process)

Equation 11 from the book is $E[X(t)] = \lambda t E[Y]$, which we are asked to prove by using conditional expectations. Since a compound Poisson process $X(t)$ is defined as $X(t) = \sum_{i=1}^{N(t)} Y_i$, where $N(t)$ is a Poisson random process, we can compute the expectation of $X(t)$ by conditioning on the value of $N(t)$, as follows. Notationally we have $E[X(t)] = E[E[X(t)|N]]$. Now this inner expectation is given by

$$
\begin{aligned}
E[X(t)|N(t)] &= E[\sum_{i=1}^{N(t)} Y_i | N(t) = n] \\
&= E[\sum_{i=1}^{n} Y_i | N(t) = n] \\
&= nE[Y_i] = nE[Y] .
\end{aligned}
$$

Thus the total expectation of $X(t)$ is given by taking the expectation of the above expression with respect to N, giving

$$
E[X(t)] = E[NE[Y]] = E[Y]E[N] = \lambda t E[Y] ,
$$

since for a Poisson process $E[N] = \lambda t$. To prove Equation 12 from the book we recall the conditional variance formula which in terms of the random variables for this problem is given by

$$\mathrm{Var}(X) = E[\mathrm{Var}(X|N)] + \mathrm{Var}(E[X|N]) .$$

From the earlier part of this problem we know that $E[X|N] = NE[Y]$, so the second term in the above expression is simply given by

$$\mathrm{Var}(E[X|N]) = \mathrm{Var}(NE[Y]) = E[Y]^2\mathrm{Var}(N) = \lambda t E[Y]^2 .$$

Where we have used the variance of a Poisson distribution ($\mathrm{Var}(N(t)) = \lambda t$). To complete this derivation we will now compute the first term in the conditional variance formula above. We begin with the expression inside the expectation, i.e. $\mathrm{Var}(X|N)$. We find that

$$\begin{aligned}
\mathrm{Var}(X|N) &= \mathrm{Var}\left(\sum_{i=1}^{N(t)} Y_i|N\right) \\
&= \mathrm{Var}\left(\sum_{i=1}^{n} Y_i|N = n\right) \\
&= \sum_{i=1}^{n} \mathrm{Var}\left(Y_i\right) \\
&= n\mathrm{Var}(Y) .
\end{aligned}$$

Where in the above we have used the fact that since the Y_i are independently identically distributed random variables the variance of the sum is the sum of the variances. Using this result the expectation of this with respect to N, and the first term in our conditional variance formula, is given by

$$E[\mathrm{Var}(X|N)] = E[N\mathrm{Var}(Y)] = \mathrm{Var}(Y)E[N] = \mathrm{Var}(Y)\lambda t .$$

Using the known result for the expectation of a Poisson process. Combining everything we find that

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}(Y)\lambda t + E[Y]^2\lambda t \\
&= (E[Y^2] - E[Y]^2)\lambda t + E[Y]^2\lambda t \\
&= E[Y^2]\lambda t ,
\end{aligned}$$

as expected.

Problem 12 (a nonhomogenous Poisson process)

We want to prove that for an nonhomogenous Poisson process with intensity function $\lambda(t)$ that

$$P\{N(t) = n\} = e^{-m(t)}\frac{m(t)^n}{n!} \quad \text{for} \quad n \geq 0 .$$

where $m(t) = \int_0^t \lambda(\tau)d\tau$. To do this we will be able to follow exactly the steps in the proof of Theorem 2.1 highlighting the differences between that situation and this one. We begin by defining $P_n(t) = P\{N(t) = n\}$, and consider $P_0(t+h)$, we find that

$$\begin{aligned} P_0(t+h) &= P\{N(t+h) = 0\} \\ &= P\{N(t) = 0, N(t+h) - N(t) = 0\} \\ &= P\{N(t) = 0\}P\{N(t+h) - N(t) = 0\}\,, \end{aligned}$$

where the last step uses the independent increment property of the nonhomogenous Poisson process. Now to evaluate $P\{N(t+h) - N(t) = 0\}$, we will use the fact that probabilities are normalized (and must sum to one) We find that

$$1 = P\{N(t+h) - N(t) = 0\} + P\{N(t+h) - N(t) = 1\} + P\{N(t+h) - N(t) \geq 2\}\,.$$

Using the assumed infinitesimal probability properties for a nonhomogenous Poisson process, we see that

$$\begin{aligned} P\{N(t+h) - N(t) = 0\} &= 1 - \lambda(t)h - o(h) - o(h) \\ &= 1 - \lambda(t)h + o(h)\,. \end{aligned}$$

This expression will be used twice in this problem. Using it once here the equation for $P_0(t+h)$ is given by

$$P_0(t+h) = P_0(t)(1 - \lambda(t)h + o(h))\,,$$

so that an approximation to the first derivative of $P_0(t)$ is given by

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda(t)P_0(t) + \frac{o(h)}{h}\,.$$

Taking the limit of both sides of this expression as $h \to 0$, we get $P_0'(t) = -\lambda(t)P_0(t)$, or

$$P_0'(t) + \lambda(t)P_0(t) = 0\,.$$

Which is different from the constant homogenous Poisson process due to the time dependence of λ in the above equation. To solve this differential equation we will introduce an integrating factor which will enable us to integrate this expression easily. Define the function $m(t)$ and $v(t)$ as

$$\begin{aligned} m(t) &\equiv \int_0^t \lambda(\tau)d\tau \\ v(t) &\equiv e^{\int_0^t \lambda(\tau)d\tau} = e^{m(t)}\,. \end{aligned}$$

This $v(t)$ is an integrating factor for the above differential equation. Multiplying the differential equation for $P_0(t)$ by $v(t)$ we see that we obtain

$$v(t)P_0'(t) + \lambda(t)v(t)P_0(t) = 0\,.$$

That v is an integrating factor of this differential equation, can be seen by noting that the derivative of v has the property that

$$v'(t) = e^{m(t)}m'(t) = e^{m(t)}\lambda(t) = v(t)\lambda(t)\,,$$

so that the v multiplied differential equation above becomes

$$v(t)P_0'(t) + v'(t)P_0(t) = 0\,.$$

or

$$\frac{d}{dt}(v(t)P_0(t)) = 0\,,$$

which can be easily integrated to give $v(t)P_0(t) = C$, for some constant C. Solving for $P_0(t)$ gives in terms of $m(t)$,

$$P_0(t) = Ce^{-m(t)}\,.$$

At the initial time $t = 0$, we are told that $N(0) = 0$ and from the definition of m we see that $m(0) = 0$. Since $P_0(0) = 1$, the constant C is seen to equal one. Thus we have found for $P_0(t)$ the following expression

$$P_0(t) = e^{-m(t)}\,.$$

This verifies the correctness of the result we are trying to prove in the case when $n = 0$. To generalize this result to higher n lets define $P_n(t) = P\{N(t) = n\}$ for $n \geq 1$. Then following similar manipulations as before we find that

$$\begin{aligned}
P_n(t+h) &= P\{N(t+h) = n\} \\
&= P\{N(t) = n, N(t+h) - N(t) = 0\} \\
&+ P\{N(t) = n-1, N(t+h) - N(t) = 1\} \\
&+ \sum_{k=2}^{n} P\{N(t) = n-k, N(t+h) - N(t) = k\} \\
&= P\{N(t) = n\}P\{N(t+h) - N(t) = 0\} \\
&+ P\{N(t) = n-1\}P\{N(t+h) - N(t) = 1\} \\
&+ \sum_{k=2}^{n} P\{N(t) = n-k\}P\{N(t+h) - N(t) = k\} \\
&= P_n(t)(1 - \lambda(t)h + o(h)) + P_{n-1}(t)(\lambda(t)h + o(h)) + o(h)\,,
\end{aligned}$$

Which is derived using independent increments and the expression for $P\{N(t+h)-N(t) = 0\}$ derived earlier. Manipulating this expression into an approximation to the first derivative of $P_n(t)$ we have that

$$\frac{P_n(t+h) - P_n(t)}{h} = -\lambda(t)P_n(t) + \lambda(t)P_{n-1}(t) + \frac{o(h)}{h}\,,$$

so that when $h \to 0$, the above approximation becomes the following differential equation

$$P_n'(t) + \lambda(t)P_n(t) = \lambda(t)P_{n-1}(t)\,.$$

Using the same integrating factor $v(t)$ introduced above (we multiply both sides by $v(t)$ and recognize the derivative of v as one of the factors) this differential equation can be seen equivalent to

$$\frac{d}{dt}(e^{m(t)}P_n(t)) = \lambda(t)e^{m(t)}P_{n-1}(t)\,, \tag{1}$$

Since we know an expression for $P_0(t)$, we can take $n = 1$ in the above to get the equation for $P_1(t)$. We find that

$$\frac{d}{dt}\left(e^{m(t)}P_1(t)\right) = \lambda(t)e^{m(t)}e^{-m(t)} = \lambda(t)\,.$$

Integrating both side of this expression we find that

$$e^{m(t)}P_1(t) = (m(t)+C)e^{-m(t)}\,,$$

for some constant C. Solving for $P_1(t)$, we have $P_1(t) = (m(t)+C)e^{-m(t)}$. To evaluate this constant C we note that since $P_1(0) = 0$, the constant C must be taken as zero. Thus

$$P_1(t) = m(t)e^{-m(t)}\,.$$

This verifies the result we are trying to prove in the case when $n = 1$. To finally complete this exercise we will prove by induction that the expression for $P_n(t)$ is true for all $n \geq 0$. Lets assume that the functional expression for $P_n(t)$ given by

$$P_n(t) = \frac{m(t)^n e^{-m(t)}}{n!}\,,$$

is valid up to some index $n-1$, and then show using the differential equation above that this formula gives the correct expression for the index n. We have already shown that this formula is true for the cases $n = 0$ and $n = 1$. To show that it is true for the index n consider Eq. 1 and insert the assumed true expression $P_{n-1}(t)$ into the right hand side. We obtain

$$\begin{aligned}\frac{d}{dt}(e^{m(t)}P_n(t)) &= \lambda(t)e^{m(t)}\left(\frac{m(t)^{n-1}e^{-m(t)}}{(n-1)!}\right)\\ &= \lambda(t)\frac{m^{n-1}(t)}{(n-1)!}\,.\end{aligned}$$

When we integrate both sides of this expression we obtain

$$e^{m(t)}P_n(t) = \int_0^t \frac{m^{n-1}(\tau)}{(n-1)!}\lambda(\tau)d\tau + C\,.$$

To evaluate this integral we let $u = m(\tau)$ so that $du = m'(\tau)d\tau = \lambda(\tau)d\tau$. With this substitution we find that our integral is given by

$$\int_{m(0)}^{m(t)} \frac{u^{n-1}}{(n-1)!}du = \frac{m(t)^n}{n!} - \frac{m(0)^n}{n!} = \frac{m(t)^n}{n!}\,,$$

since $m(0) = 0$. Thus the expression for $P_n(t)$ above becomes

$$P_n(t) = \left(\frac{m(t)^n}{n!} + C\right)e^{-m(t)}\,.$$

Evaluate this function at $t = 0$ since $P_n(0) = 0$, we see that the constant C must be equal to zero. Thus we have shown that for all $n \geq 0$ that

$$P\{N(t) = n\} = \frac{m(t)^n e^{-m(t)}}{n!}\,,$$

as requested.

Problem 13 (the distribution of the event times in a nonhomogenous process)

Following the same strategy that the book used to compute the distribution function for a homogenous Poisson process, we will begin by assuming an ordered sequence of arrival times

$$0 < t_1 < t_2 < \cdots < t_n < t_{n+1} = t$$

and let h_i be small increments such that $t_i + h_i < t_{i+1}$, for $i = 1, 2, \cdots n$. Then the probability that a random sample of n arrival times S_i happen at the times t_i and conditioned on the fact that we have n arrivals by time t can be computed by considering

$$P\{t_i \le S_i \le t_i + h_i, \quad \text{for} \quad i = 1, 2, \cdots, n | N(t) = n\}.$$

Which if we define the event A to be the event that we have exactly *one* event in $[t_i, t_i + h_i]$ for $i = 1, 2, \cdots, n$ and *no* events in the other regions then (by definition) the above equals the following expression

$$\frac{P\{A\}}{P\{N(t) = n\}}.$$

The probability that we have one event in $[t_i, t_i + h_i]$ is given by the fourth property in the definition of a nonhomogenous Poisson and is given by

$$P\{N(t_i + h_i) - N(t_i) = 1\} = \lambda(t_i)h_i + o(h_i) \tag{2}$$

To calculate the probability that we have *no* events in a given interval, we will derive this from the four properties in the definition a nonhomogenous Poisson process. Specifically, since the total probability must sum to one we have the constraint on increment variables over the range of time $[t_l, t_r]$,

$$P\{N(t_r) - N(t_l) = 0\} + P\{N(t_r) - N(t_l) = 1\} + P\{N(t_r) - N(t_l) \ge 2\} = 1.$$

or using properties (ii) and (iii) in the definition of a nonhomogenous Poisson process the above becomes (solving for $P\{N(t_r) - N(t_l) = 0\}$) the following

$$P\{N(t_r) - N(t_l) = 0\} = 1 - \lambda(t_l)(t_r - t_l) + o(t_r - t_l). \tag{3}$$

This result will be used in what follows. To evaluate $P\{A\}$ we recognized that in the intervals

$$(0, t_1), (t_1 + h_1, t_2), (t_2 + h_2, t_3), \cdots, (t_n + h_n, t_{n+1}),$$

no events occurs, while in the intervals

$$(t_1, t_1 + h_1), (t_2, t_2 + h_2), (t_3, t_3 + h_3), \cdots, (t_n, t_n + h_n),$$

one event occurs. By the independent increments property of nonhomogenous process the event A can be computed as the product of the probabilities of each of the above intervals event. The contributed probability $P(A_1)$ in the evaluation of $P\{A\}$ from the intervals where the count increase by *one* is given by

$$P(A_1) = \prod_{i=1}^{n} \{\lambda(t_i)h_i + o(h_i)\} = \prod_{i=1}^{n} \lambda(t_i)h_i + o(h),$$

where we have used Eq. 2 and the term $o(h)$ represents terms higher than first order in any of the h_i's. By analogy with this result the contributed probability in the evaluation of $P\{A\}$ from the intervals where the count does not increase $P(A_0)$ is given by

$$P(A_0) = (1 - \lambda(0)(t_1) + o(t_1)) \prod_{i=1}^{n} \{1 - \lambda(t_i + h_i)(t_{i+1} - t_i - h_i) + o(t_{i+1} - t_i - h_i)\} .$$

This expression will take some manipulations to produce a desired expression. We begin our sequence of manipulations by following the derivation in the book and recognizing that we will eventually be taking the limits as $h_i \to 0$. Since this expression has a finite limit we can take the limit of the above expression as is and simplify some of the notation. Taking the limit $h_i \to 0$ and defining $t_0 = 0$ the above expression becomes

$$P(A_0) = \prod_{i=0}^{n} \{1 - \lambda(t_i)(t_{i+1} - t_i) + o(t_{i+1} - t_i)\} .$$

We can simplify this product further by observing that the individual linear expressions we multiply can be written as an exponential which will facilitate our evaluation of this product. Specifically, it can be shown (using Taylor series) that

$$e^{-\lambda(t_i)(t_{i+1}-t_i)} = 1 - \lambda(t_i)(t_{i+1} - t_i) + o(t_{i+1} - t_i) .$$

With this substitution the product above becomes a sum in the exponential and we have

$$P(A_0) = \prod_{i=0}^{n} e^{-\lambda(t_i)(t_{i+1}-t_i)} = \exp\left\{ - \sum_{i=0}^{n} \lambda(t_i)(t_{i+1} - t_i) \right\} .$$

Recognizing the above summation as an approximation to the integral of $\lambda(\cdot)$, we see that the above is approximately equal to the following

$$P(A_0) = \exp\left\{ - \sum_{i=0}^{n} \lambda(t_i)(t_{i+1} - t_i) \right\} \approx \exp\left\{ - \int_0^t \lambda(\tau) d\tau \right\} = e^{-m(t)} .$$

With these expressions for $P(A_1)$ and $P(A_0)$, we can now evaluate our target expression

$$\begin{aligned}
\frac{P\{A\}}{P\{N(t) = n\}} &= \frac{P(A_1)P(A_0)}{P\{N(t) = n\}} \\
&= \left(\frac{n!}{e^{-m(t)} m(t)^n} \right) P(A_1)P(A_0) \\
&= \left(\frac{n!}{e^{-m(t)} m(t)^n} \right) \left(\left(\prod_{i=1}^{n} \lambda(t_i) h_i + o(h) \right) e^{-m(t)} \right) \\
&= n! \left(\prod_{i=1}^{n} \frac{\lambda(t_i)}{m(t)} h_i + o(h) \right) .
\end{aligned}$$

It is this final result we were after. After dividing by $\prod_{i=1}^{n} h_i$ and taking the limit where $h_i \to 0$, we can conclude that the probability of drawing a specific sample of n event times (i.e. obtaining a draw of the random variables S_i) for a *nonhomogenous* Poisson process with rate $\lambda(t)$ given that we have seen n events by time t is given by

$$f_{S_1,S_2,\cdots,S_n}(t_1,t_2,\cdots,t_n|N(t)=n) = n!\left(\prod_{i=1}^{n}\frac{\lambda(t_i)}{m(t)}\right) \quad 0 < t_1 < t_2 < \cdots < t_n < t \tag{4}$$

We recognized that this expression is the same distribution as would be obtained for the order statistics corresponding to n independent random variables uniformly distributed with probability density function $f(\cdot)$ and a cumulative distribution function $F(\cdot)$ given by

$$f(x) = \frac{\lambda(x)}{m(t)} \quad \text{and} \quad F'(x) = f(x)\,.$$

By the definition of the function $m(\cdot)$ we have that $\lambda(x) = m'(x)$, so that an equation for our cumulative distribution function F is given by

$$F'(x) = \frac{m'(x)}{m(t)}\,.$$

This can be integrated to give

$$F(x) = \frac{m(x)}{m(t)}\,,$$

which can only hold if $x \leq t$, while if $x > t$, $F(x) = 1$. This is the desired result.

References

[1] S. Ross. *A First Course in Probability.* Macmillan, 3rd edition, 1988.