Homework 3 – February 15, 2006 **Solution** prepared by Tobin Fricke
Physics 402 – Probability Prof. S.G. Rajeev, Univ. of Rochester

7.1. Waiting times in a Poisson Process *A Geiger counter emits a click each time a radioactive decay happens. If the average number of decays in unit time is* λ*, what is the proability distribution of the time interval between clicks?*

The decay of a bulk quantity of radioactive material is one example of a Poisson process. For a Poisson process, we expect the number of events occuring in a unit time interval to follow the Poisson distribution, $p[k \text{ events in unit time}] = e^{\lambda}\lambda^k/k!$, where λ is the expected number of events in the unit time interval. This must old for any notion of a "unit time interval," so we may interpret λ as the *rate* of events. The distribution of number of events occurring in an arbitrary time duration t is therefore also Poisson, with parameter λt. The probability of receiving exactly k events in a time interval t is therefore $p[N_t = k] = \exp(-\lambda t)(\lambda t)^k/k!$.

The condition of having a waiting time T until the first event is equivalent to having exactly zero events in time N_T followed by exactly one event in the following time dt. For a poisson process, the probability of getting an event in an interval $[t, t+dt]$ is λdt. Consequently the probability distribution of waiting times is $p[N_t = 0]\lambda dt = \exp(-\lambda t)\lambda dt$. The probability density of waiting times is therefore $\boxed{\lambda \exp(-\lambda t)}$.

Note that this integrates to unity over the range $t \in (0, \infty)$.

We may also solve for the probability density of waiting times by writing and then solving an integral relation.

Let $p(t)dt$ be the probability that the waiting time is in $[t, t+dt]$. Then $\int_0^t p(t')dt'$ is the probability that the waiting time is less than t. Subtract this from unity to get the probability that the waiting time is at least t, $1 - \int_0^t p(t')dt'$. Consider the probability that the waiting time is at least t and an event happens in the following time interval dt; these are independent events, so we may just multiply by the probability λdt of an event occurring in a duration dt. We've recovered an expression for the probability that the waiting time is between t and $t + dt$, giving us the integral relation:

$$\left(1 - \int_0^t p(t')dt'\right) \lambda \cancel{dt} = p(t)\cancel{dt}$$

We can begin to solve this by taking the derivative with respect to t. Note that $\int_0^t f(t')dt' = F(t) - F(0)$, where $F'(t) = f(t)$, so $(d/dt)\int_0^t f(t')dt' = f(t)$. We get:

$$p'(t) = \frac{d}{dt}\left(1 - \int_0^t p(t')dt'\right)\lambda = -\lambda p(t)$$

This has the well-known solution

$$p(t) = Ce^{-\lambda t}$$

where C is some constant. The normalization requirement $\int_0^\infty p(t)dt = 1$ gives us

$$\boxed{p(t) = \lambda e^{-\lambda t}}$$

which agrees with what we found earlier.

7.2. Nearest neighbor distances between randomly spaced points *Assume that homes in thep rarie are distributed uinformly with an average density of* n *per square mile. What is the probability distribution of the distance to the nearest neighbor from a given home? What is the average distance between nearest neighbors?*

This question is similar to 7.1 above; instead of a “waiting time” until the next event after some arbitrary starting time, we’re interested in the “waiting distance” as we travel radially outward from a given point until we encounter another house.

Approaching this using the integral relation technique, we may write

$$\left(1 - \int_0^r p(r')dr'\right) 2\pi r n dr = p(r)dr$$

where the density n fills the role of λ in the one-dimensional case.

Requiring that $p(0) = 0$ and $\int_0^\infty p(r)dr = 1$ (note the lower limit of integration; the probability of a negative waiting time is zero), we find

$$\boxed{p(r) = 2\pi n r \exp\{-n\pi r^2\}}$$

The average nearest-neighbor distance is given by

$$\boxed{\langle r_{\text{nearest neighbor}}\rangle = \int_0^\infty r p(r)dr = \frac{1}{2\sqrt{n}}}$$

8. Transformation of Random Variables

There is a nice explanation of this in *Numerical Recipes in C*, chapter 7.2. The text is available freely at `http://www.library.cornell.edu/nr/bookcpdf/c7-2.pdf`.

We employ conservation of probability: $|p_y(y)dy| = |p_x(x)dx|$. We are given that the probability density of x is uniform, i.e. $p_x(x) = 1$, and we are given several desired probability densities p_y. The procedure is to solve for the derivative dx/dy, obtain x in terms of y by integrating, and then invert the relation to get y in terms of x.

$$|p_y(y)dy| = |p_x(x)dx|$$

$$\frac{dx}{dy} = \pm p_y(y)$$

$$x = \int \frac{dx}{dy} dy = \pm \int p_y(y)dy$$

The following Mathematica code peforms this procedure for $p_y(y) = -e^{-y}$:

```
py[y_] := -Exp[-y]
Solve[x==Integrate[py[y],y],y]
```

We find that $y_1(x) = -\log(x)$, $y_2(x) = \text{erf}^{-1}(2x - 1)$, and $y_3(x) = \tan(\pi x)$.

9. Multiplicative Random Walk *Consider the following simple model for the size of a colony of bacteria. We start with a number n_0; in each generation the number can be either multiplied by a factor u with probability $1/2$ or divided by the same number with the same probability. What is the probability distribution of the number of bacteria after a large number N of steps? The number u is near unity.*

If η is the random variable giving the population of the colony after many steps,then we may write η as a product over many random variables η_i each of which may attain the values u and $1/u$, describing how the size of the colony changes in the ith step:

$$\eta = \eta_1 \eta_2 \eta_3 \cdots \eta_N$$

If we take the logarithm, then the product is converted into a sum:

$$\log\eta = \log\eta_1 + \log\eta_2 + \log\eta_3 + \cdots \log|eta_N$$

The central limit theorem tells us that the sum of many independent random variables will have a normal (Gaussian) distribution. If the logarithm of η follows the normal distribution, then η itself follows the so-called log-normal distribution (see `http://en.wikipedia.org/wiki/Log_normal`).

One may also write the final size of the colony as $y = n_0 u^x (1/u)^{N-k} = n_0 u^{2x-N}$, where x is the random variable giving the number of successes in N Bernoulli trials; k will in general follow the binomial distribution with mean $\mu = Np = N/2$ and variance $\sigma^2 = npq = N/4$, but for large N this converges to the normal distribution with the same mean and variance.

Use the conservation of probability formula:

$$|p(y)dy| = |p(x)dx|$$

to get:

$$p(y) = \frac{dx}{dy}p(x)$$

Solve the expression $y = n_0 u^{2x-N}$ for x:

$$x = \frac{\log(y/n_0)}{2\log u} + \frac{N}{2}$$

Take the derivative:

$$\frac{dx}{dy} = \frac{1}{2y\log u}$$

We know that $p(x)$ is the probability density of the normal distribution with mean $\mu = N/2$ and variance $\sigma^2 = N/4$:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} = \sqrt{\frac{2}{\pi N}} \exp\left\{-\frac{(N-2x)^2}{2N}\right\} = \sqrt{\frac{2}{\pi N}} \exp\left\{-\frac{-\left(\log\frac{y}{n_0}\right)^2}{2N(\log u)^2}\right\}$$

So we have $p(y) = \frac{dx}{dy}p(x)$, which becomes, with everything plugged in:

$$\boxed{p(y) = \frac{1}{y(\log u)\sqrt{2\pi N}} \exp\left\{-\frac{-\left(\log\frac{y}{n_0}\right)^2}{2N(\log u)^2}\right\}}$$

Note that this is the probability density function of the log-normal distribution with mean $\mu = \log(n_0)$ and variance $\sigma^2 = N(\log u)^2$.