**PHY402 Probability Spring 2006**

**Final Exam Wednesday March 22 2006 9:30-10:45am**

**1** If $\xi_1$ and $\xi_2$ are independent uniform random variables in the range $[0, 1]$ what are the probability distributions of the sum $\xi_1 + \xi_2$ and product $\xi_1 \xi_2$?

**2** A random variable is known to take values in the range $[a, b]$ with a mean of $\mu$. What is the probability distribution of maximum entropy given only these facts? Find the the entropy of this distribution explicitly. Which value of $\mu$ (for given $a, b$) gives the largest entropy even among these distributions?

**Solutions**

1. The the probability density function of $\eta = \xi_1 + \xi_2$ is

$$p_\eta(y) = \int_0^1 \int_0^1 \delta(x_1 + x_2 - y) p_{\xi_1}(x_1) p_{\xi_2}(x_2) dx_1 dx_2 \tag{1}$$

which is also the convolution

$$p_\eta(y) = \int_0^1 p_{\xi_1}(x_1) p_{\xi_2}(y - x_1) dx_1 \tag{2}$$

Since $\xi_1, \xi_2$ are uniform in the range $[0, 1]$, their p.d.f. is equal to one in $[0, 1]$ and zero otherwise [1] :

$$p_\eta(y) = \int_0^1 \theta(0 < y - x_1 < 1) dx_1 \tag{3}$$

The range of $\eta$ is $[0, 2]$. If it takes a value $y < 1$ less than the midpoint, the condition $y - x_1 < 1$ is automatically satisfied since $x_1 > 0$. So we get

$$p_\eta(y) = \int_0^1 \theta(0 < y - x_1) dx_1 = \int_0^y dx_1. \tag{4}$$

Thus

$$p_\eta(y) = y, \quad \text{for } y < 1. \tag{5}$$

If on the other hand $y > 1$, the condition $0 < y - x_1$ is automatic since $x_1 < 1$. Then

$$p_\eta(y) = \int_0^1 \theta(y - x_1 < 1) dx_1 = \int_{y-1}^1 dx_1 \tag{6}$$

so that

$$p_\eta(y) = 2 - y, \quad \text{for } y > 1. \tag{7}$$

The two expressions coincide for $y = 1$. There is also a method using characterstic functions but this is harder.

For the product $\zeta = \xi_1 \xi_2$

$$p_\zeta(z) = \int p_{\xi_1}(x_1) p_{\xi_2}\left(\frac{z}{x_1}\right) \frac{dx_1}{x_1} \tag{8}$$

which follows from

$$p_\zeta(z) = \int_0^1 \int_0^1 \delta(x_1 x_2 - z) p_{\xi_1}(x_1) p_{\xi_2}(x_2) dx_1 dx_2 \tag{9}$$

[1] $\theta(0 < x < 1)$ is the function that is equal to one in the range $[0, 1]$ and zero otherwise

Thus

$$p_\zeta(z) = \int_0^1 \theta(0 < \frac{z}{x_1} < 1)\frac{dx_1}{x_1} = \int_z^1 \frac{dx_1}{x_1} = -\log z. \tag{10}$$

since the condition $0 < \frac{z}{x_1}$ is automatically satisfied. Thus

$$p_\eta(z) = -\log z. \tag{11}$$

Another method is to note that $\log \zeta = \log \xi_1 + \log \xi_2$ and that $\log \xi_1, \log \xi_2$. Then we must determine the distributions of $\log \xi_1, \log \xi_2$ (change of variables) then evaluate the convolution as above. The methods using characteristic functions are harder in this case.

2. We need to maximize the entropy

$$-\int p(x) \log p(x) dx \tag{12}$$

subject to the conditions

$$\int_a^b p(x)dx = 1, \quad \int_a^b xp(x)dx = \mu \tag{13}$$

and the inequalities

$$p(x) > 0, \quad p(x) = 0 \text{ unless } x \in [a, b]. \tag{14}$$

Using Lagrange multipliers we must maximize

$$F(p) = -\int p(x) \log p(x) dx + \lambda_0 \int p(x)dx + \lambda_1 \int xp(x)dx. \tag{15}$$

This gives

$$\delta F = \int \delta p(x)[-\log p(x) - 1 + \lambda_0 + \lambda_1 x]dx = 0. \tag{16}$$

The solution is the exponential distribution

$$p(x) = e^{\lambda_0 - 1} e^{\lambda_1 x}, \text{for } x \in [a, b] \tag{17}$$

and zero outside this range. The constants are determined from the constraints:

$$\int_a^b p(x)dx = e^{\lambda_0 - 1}\left[\frac{e^{\lambda_1 b} - e^{\lambda_1 a}}{\lambda_1}\right] = 1, \tag{18}$$

Thus

$$p(x) = \frac{1}{Z(\lambda_1)} e^{\lambda_1 x}, \text{for } x \in [a, b] \tag{19}$$

where

$$Z(\lambda_1) = \left[\frac{e^{\lambda_1 b} - e^{\lambda_1 a}}{\lambda_1}\right] = \int_a^b e^{\lambda_1 x} dx. \tag{20}$$

To determine $\mu$ we note that

$$\mu = \int_a^b x p(x) dx = \frac{1}{Z(\lambda_1)} \frac{\partial}{\partial \lambda_1} \int_a^b e^{\lambda_1 x} dx. \tag{21}$$

which gives us the mean as a function of $\lambda$:

$$\frac{\partial}{\partial \lambda_1} \log Z(\lambda_1) = \mu. \tag{22}$$

That is

$$-\frac{1}{\lambda_1} + \frac{b e^{\lambda_1 b} - a e^{\lambda_1 a}}{e^{\lambda_1 b} - e^{\lambda_1 a}} = \mu \tag{23}$$

1.It won't be necessary to invert this formula for what follows. The trick is to think of $\lambda_1$ as the independent variable instead of $\mu$

2.You can also just evaluate $\int_a^b x e^{\lambda_1 x} dx$ by integrating by parts, but I use here a method that would work in more general situations as well.

3.Note for later use that as $\lambda_1 \to 0$, (uniform distribution)

$$\mu \to \frac{a+b}{2}. \tag{24}$$

The entropy is

$$S = -\int p(x) \log p(x) dx = < -\log p(x) > = -\lambda_1 < x > + \log Z(\lambda_1) = -\lambda_1 \mu + \log Z(\lambda_1). \tag{25}$$

The condition for maximizing entropy is

$$\frac{\partial}{\partial \lambda_1} S = 0. \tag{26}$$

Using the chain rule of calculus,

$$\frac{\partial}{\partial \lambda_1} S = -\mu - \lambda_1 \frac{\partial \mu}{\partial \lambda_1} + \mu = -\lambda_1 \frac{\partial \mu}{\partial \lambda_1} \tag{27}$$

Thus the maximum occurs either when $\lambda_1 = 0$ or when $\frac{\partial \mu}{\partial \lambda_1} = 0$. Let us see what the simple solution $\lambda_1 = 0$ means in our problem.

4.Try harder things only if the easy route does not work. In fact $\frac{\partial \mu}{\partial \lambda_1} > 0$ so it cannot vanish: this quantity is actually the variance. But you don't need to know that to see that $\lambda_1 = 0$ is a perfectly good solution.

When $\lambda_1 = 0$, we get the uniform distribution.

$$p_\xi(x) = \frac{1}{b-a}, \qquad \text{for} x \in [a, b]. \tag{28}$$

The value of the mean is then just the midpoint $\mu = \frac{a+b}{2}$.

5.It is not hard to expand $Z$ and hence $S$ as a function of $\lambda_1$ around $\lambda_1 = 0$ to see that this extremum is in fact a maximum.

$$Z(\lambda_1) \approx (b-a)[1 + \lambda_1 \frac{a+b}{2} + \frac{\lambda_1^2}{3!} \frac{(b^3 - a^3)}{b-a} + \cdots] \tag{29}$$

etc. This is more than you need to do solve this problem.