Stochastic Optimisation

Solutions to Problem Sheet 3

---

1. One way to do this is by computing the complementary cdf,

$$1 - F_X(x) = \mathbb{P}(X > x) = \mathbb{P}((-\log U)/\lambda > x) = \mathbb{P}(\log U < -\lambda x)$$
$$= \mathbb{P}(U < e^{-\lambda x}) = \begin{cases} 1, & x \le 0 \\ e^{-\lambda x}, & x > 0. \end{cases}$$

This matches the ccdf of an Exp(λ) random variable.

Another way is to use the change of variables formula. Define the function $g : [0,1] \to \mathbb{R}_+$ by $g(u) = (-\log u)/\lambda$. Then, $x = g(u)$ if and only if $u = e^{-\lambda x}$, i.e., $g^{-1}(x) = e^{-\lambda x}$, where g^{-1} maps $\mathbb{R}_+$ to $[0,1]$. Moreover, $|g'(u)| = 1/(\lambda u)$, so that $|g'(g^{-1}(x)| = e^{\lambda x}/\lambda$. The density of the random variable U is $f_U(u) = 1$ for $u \in [0,1]$, and zero outside this interval. We now have by the change of variables formula that

$$f_X(x) = \frac{1}{|g'(g^{-1}(x)|} f_U(g^{-1}(x)) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

This matches the density of an Exp(λ) random variable.

2. This can be shown quite easily using the change of variables formula. An even easier way is to use the relation between Beta and Gamma random variables. If X has a Beta(α, β) distribution, then we can write $X = \frac{V}{V+W}$, where V and W are independent and have Gamma(α,1) and Gamma(β,1) distributions. Now,

$$Y = 1 - X = \frac{W}{W+V},$$

and so it must have a Beta(β, α) distribution.

3. This is a straightforward, but somewhat long, calculation using the formula for transformations of random variables. Define $g : \mathbb{R}_+ \times [0, 2\pi) \to \mathbb{R}^2$ by $g(x,\theta) = (\sqrt{x}\sin\theta, \sqrt{x}\cos\theta)$. Then, $(V,W) = g(X,\Theta)$. Now, X has density $f_X(x) = \frac{1}{2}\exp(-\frac{x}{2})$ on $\mathbb{R}_+$, and Θ has density $f_\Theta(\theta) = \frac{1}{2\pi}$ on $[0,2\pi)$, and these are independent random variables. Hence, their joint density is

$$f_{X,\Theta}(x,\theta) = \frac{1}{4\pi}e^{-x/2}, \quad (x,\theta) \in \mathbb{R}_+ \times [0,2\pi),$$

and zero outside this set.

Next, we compute the Jacobian matrix of g and its determinant. We have,

$$J_g(x,\theta) = \begin{pmatrix} \frac{\sin\theta}{2\sqrt{x}} & \sqrt{x}\cos\theta \\ \frac{\cos\theta}{2\sqrt{x}} & -\sqrt{x}\sin\theta \end{pmatrix}, \quad |\det J_g(x,\theta)| = \Big| \frac{-(\sin^2\theta + \cos^2\theta)}{2} \Big| = \frac{1}{2}.$$

Moreover, if $g(x,\theta) = (v,w)$, then (x,θ) are uniquely determined, and $x = v^2 + w^2$. Hence, we get

$$f_{V,W}(v,w) = \sum_{(x,\theta):(v,w)=g(x,\theta)} \frac{1}{|\det J_g(x,\theta)|} F_{X.\Theta}(x,\theta)$$
$$= 2\frac{1}{4\pi}e^{-(v^2+w^2)/2} = \frac{1}{\sqrt{2\pi}}e^{-v^2/2}\frac{1}{\sqrt{2\pi}}e^{-w^2/2}.$$

Thus, the joint density of (V, W) factorises into the product of standard Gaussian densities for V and W, which implies that V and W are independent, standard Gaussian random variables.

4. As stated in the question, the posterior density for θ, given that we observe x, is proportional to the product of the prior density, and the likelihood of observing x when the parameter value is θ, i.e.,

$$\pi_1(\theta|x) \propto \pi_0(\theta) f_\theta(x).$$

Now π_0 is the density of a $N(\mu_0, \sigma_0^2)$ random variable, so

$$\pi_0(\theta) \propto \exp\Bigl(-\frac{(\theta-\mu_0)^2}{2\sigma_0^2}\Bigr),$$

whereas x has an $N(\theta, 1)$ density, i.e.,

$$f_\theta(x) \propto \exp\Bigl(-\frac{(x-\theta)^2}{2}\Bigr).$$

Putting these together, we obtain that

$$\begin{aligned}
\pi_1(\theta|x) &\propto \exp\Bigl(-\frac{(\theta-\mu_0)^2}{2\sigma_0^2} - \frac{(x-\theta)^2}{2}\Bigr) \\
&\propto \exp\Bigl(-\frac{1}{2\sigma_0^2}\Bigl[\theta^2 - 2\mu_0\theta + \mu_0^2 + \sigma_0^2\bigl(x^2 - 2x\theta + \theta^2\bigr)\Bigr]\Bigr) \\
&\propto \exp\Bigl(-\frac{1}{2\sigma_0^2}\bigl[\theta^2 - 2\mu_0\theta + \sigma_0^2\theta^2 - 2x\sigma_0^2\theta\bigr]\Bigr) \\
&\propto \exp\Bigl(-\frac{1+\sigma_0^2}{2\sigma_0^2}\Bigl[\theta^2 - \frac{2(\mu_0 + x\sigma_0^2)\theta}{1+\sigma_0^2}\Bigr]\Bigr) \\
&\propto \exp\Bigl(-\frac{1+\sigma_0^2}{2\sigma_0^2}\Bigl[\theta - \frac{\mu_0 + x\sigma_0^2}{1+\sigma_0^2}\Bigr]^2\Bigr).
\end{aligned}$$

The first line is obtained by plugging in the expressions for π_0 and f_θ. The second line is expanding the squares. The third line drops some terms that don't depend on θ, which can be absorbed into the constant of proportionality. The fourth line groups together the θ^2 terms and the θ terms separately, and re-arranges them. The last line completes the square, and ignores the constant term that this produces. The different shapes of brackets are just for improving readability.

The last expression above, is up to constant terms, the density function of a normal random variable with mean μ_1 and variance σ_1^2, where

$$\mu_1 = \frac{\mu_0 + x\sigma_0^2}{1+\sigma_0^2}, \quad \sigma_1^2 = \frac{\sigma_0^2}{1+\sigma_0^2}.$$

5. The Thompson sampling algorithm for arms with normally distributed rewards, with known variances and unknown means, is as follows. We will assume without loss of generality that the variances are unity, as in the question; if the variances are known, the rewards can always be rescaled to have unit variance, and the algorithm modified suitably to take that into account.

   (a) Start with independent priors $\pi_{1,0}$ and $\pi_{2,0}$ for the unknown mean rewards μ_1 and μ_2 from the two arms. Take $\pi_{j,0}$ to have a $N(\mu_{j,0}^2, \sigma_{j,0}^2)$ density, $j = 1, 2$. The choice of the parameters $\mu_{j,0}$ and $\sigma_{j,0}^2$ is arbitrary, except that $\sigma_{j,0}$ shouldn't be zero. If there is no prior knowledge to guide the choice, then one possibility is to take $\mu_{j,0} = 0$ and $\sigma_{j,0}^2 = 1$ for $j = 1$ and 2.

   Broadly speaking, $\mu_{j,0}$ should be close to what you think the true mean is, and $\sigma_{j,0}$ should be about as large as your uncertainty about $\mu_{j,0}$. For example, if you think the true mean is between 100 and 200, then $\mu_{j,0} = 150$, $\sigma_{j,0} = 50$ is a reasonable choice.

(b) In time step $t+1$, sample $\theta_1(t)$ from $\pi_{1,t}$ and $\theta_2(t)$ independently from $\pi_{2,t}$. Play the arm corresponding to whichever of these samples is larger.

(c) Update the posterior for the arm which is played, based on the observed reward. Thus, if arm j is played and reward x obtained, then $\pi_{j,t+1}$ has a normal distribution, $N(\mu_{j,t+1}, \sigma^2_{j,t+1})$, where

$$\mu_{j,t+1} = \frac{\mu_{j,t} + x\sigma^2_{j,t}}{1+\sigma^2_{j,t}}, \quad \sigma^2_{j,t+1} = \frac{\sigma^2_{j,t}}{1+\sigma^2_{j,t}}.$$

The posterior for the arm that is not played remains unchanged.

(d) Increment t and go back to Step (b).