# Stochastic Optimisation - Problem Sheet 3

Dom Hutchinson

November 10, 2020

**Question 1)**
Let $U \sim \text{Unif}[0,1]$ and let $\lambda > 0$ be a given constant.
Show that the random variable $X := \frac{1}{\lambda}(-\ln U)$ has an exponential distribution with parameter $\lambda$.

**Answer 1)**

$$\begin{array}{rrcl} & F_X(x) & = & \mathbb{P}(X \leq x) \\ & & = & \mathbb{P}\left(-\frac{1}{\lambda}\ln(U) \leq x\right) \\ & & = & \mathbb{P}\left(U \geq e^{-\lambda x}\right) \\ & & = & \int_{e^{-\lambda x}}^1 1dt \\ & & = & [t]_{e^{-\lambda x}}^1 \\ & & = & 1 - e^{-\lambda x} \\ \implies & f_X(x) & = & F_X'(x) \\ & & = & \lambda e^{-\lambda x} \\ \implies & X & \sim & \text{Exp}(\lambda) \end{array}$$

**Question 2)**
Let $X$ be a random variable with a $\text{Beta}(\alpha,\beta)$ distribution and let $Y := 1 - X$.
Show that $Y$ has a $\text{Beta}(\beta,\alpha)$ distribution.

**Answer 2)**

$$\begin{array}{rcl} f_Y(y) & = & \mathbb{P}(Y = y) \\ & = & \mathbb{P}(1 - X = y) \\ & = & \mathbb{P}(X = 1 - y) \\ & = & \dfrac{(1-y)^{\alpha-1}y^{\beta-1}}{B(\alpha,\beta)} \\ \implies Y & \sim & \text{Beta}(\beta,\alpha) \end{array}$$

**Question 3)**
Let $X \sim \text{Exp}\left(\frac{1}{2}\right)$ and $\Theta \sim [0, 2\pi)$ with $X$ and $\Theta$ being independent.
Define $V := \sqrt{X}\sin\Theta$ and $W := \sqrt{X}\cos\Theta$.
Show that $V$ and $W$ are independent random variables with a $\text{Normal}(0,1)$ distribution.

**Answer 3)**
Let $X \sim \text{Exp}\left(\frac{1}{2}\right)$ and $\Theta \sim [0, 2\pi)$ with $X$ and $\Theta$ being independent.
Note that

$$f_X(x) = \frac{1}{2}e^{\frac{1}{2}x} \quad \text{and} \quad f_\Theta(\theta) = \frac{1}{2\pi}\mathbb{1}\{\theta \in [0, 2\pi)\}$$

Since $X$ and $\Theta$ are independent

$$f_{X,\Theta}(x, \theta) = f_X(x)f_\Theta(\theta) = \mathbb{1}\{\theta \in [0, 2\pi)\}\frac{1}{4\pi}e^{-\frac{1}{2}x}$$

Define random variables $V, W$ and function $g$ such that

$$(V, W) = g(X, \Theta) = \left(\sqrt{X}\sin(\Theta), \sqrt{X}\cos(\Theta)\right)$$

The Jacobian of $g$ is

$$J_g(x, \theta) = \begin{pmatrix} \frac{1}{2\sqrt{x}}\sin(\theta) & \frac{1}{2\sqrt{x}}\cos(\theta) \\ \sqrt{x}\cos(\theta) & -\sqrt{x}\sin(\theta) \end{pmatrix}$$

The determinant of $J_g$ is

$$\begin{aligned}
\det(J_g) &= (\frac{1}{2\sqrt{(x)}}\sin(\theta)) \cdot (-\sqrt{x}\sin(\theta)) - (\frac{1}{2\sqrt{x}}\cos(\theta)) \cdot (\sqrt{x}\cos\theta)) \\
&= -\frac{1}{2}\sin^2(\theta) - \frac{1}{2}\cos^2(\theta) \\
&= -\frac{1}{2}(\sin^2(\theta) + \cos^2\theta) \\
&= -\frac{1}{2}
\end{aligned}$$

For a fixed $(v, w)$ consider the set of values $(x, \theta)$ st $(v, w) = g(x, \theta)$

$$\begin{aligned}
& \{(x, \theta) : (v, w) = g(x, \theta)\} \\
\Leftrightarrow \; & \{(x, \theta) : (v, w) = (\sqrt{x}\sin(\theta), \sqrt{x}\cos(\theta))\} \\
\Leftrightarrow \; & \{(x, \theta) : x = v^2 + w^2, \theta = \arctan(v/w)\}
\end{aligned}$$

Now I derive the joint distribution of $V$ and $W$

$$\begin{aligned}
f_{V,W}(v, w) &= \sum_{\{(x,\theta):(v,w)=g(x,\theta)\}} f_{X,\Theta}(x, \theta)\frac{1}{|\det(J_g(x, \theta))|} \\
&= \sum_{\{(x,\theta):x=v^2+w^2,\theta=\arctan(v/w)\}} \mathbb{1}\{\theta \in [0, 2\pi)\}\frac{1}{4\pi}e^{-\frac{1}{2}x} \cdot \left|\frac{1}{-1/2}\right| \\
&= \frac{1}{4\pi}e^{-\frac{1}{2}(v^2+w^2)} \cdot \left|\frac{1}{-1/2}\right| \text{ since } \arctan(v/w) \in [0, 2\pi) \;\forall\; v, w \\
&= \frac{1}{2\pi}e^{-\frac{1}{2}(v^2+w^2)} \\
&= \underbrace{\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}v^2}\right)}_{\Phi} \cdot \underbrace{\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2}\right)}_{\Phi}
\end{aligned}$$

By noting that both terms in the final expression are $\Phi$ the pdf for a standard normal distribution, it is shown that $V$ and $W$ are independent standard normal random variables.

**Question 4)**
Consider a Bayesian approach to the problem of inferring the mean of a normal distribution with unknown mean and known variance. More precisely, suppose $X \sim \text{Normal}(\theta, 1)$ with $\theta$

unknown. Fix $\mu_0 \in \mathbb{R}$ and $\sigma_0^2 > 0$ and let $\pi_0$ denote the density of a Normal$(\mu_0, \sigma_0^2)$ distribution. Suppose $\pi_0$ denotes the prior distribution of $\theta$. Let $\pi_1(\cdot|x)$ denote the posterior distribution, conditional on observing a sample of $X$ which takes the value $x$. In other words

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

where $f_\theta$ denotes the density of a Normal$(\theta, 1)$ random variable. The constant of proportionality will be determined by the requirement that $\pi_1(\cdot|x)$ be the probability density (ie that it integrates to 1).

Show that $\pi_1(\cdot|x)$ is the density of a Normal$(\mu_1, \sigma_1^2)$ random variable where

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

**Answer 4)**
Let $X \sim$ Normal$(\theta, 1)$ with $\theta$ unknown and fix $\mu_0 \in \mathbb{R}, \sigma_0^2 > 0$.
Let $\pi_0 \sim$ Normal$(\mu_0, \sigma_0^2)$ be the prior for $\theta$ and $\pi_1(\cdot|x)$ be the posterior for $\theta$ given $x$ is observed from $X$. This means

$$\pi_1(\theta|x) \propto \pi_0(\theta) f_\theta(x) \quad \text{where } f_\theta(x) := \mathbb{P}(X = x|\theta)$$

Note that

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2} \quad \text{and} \quad \pi_0(\theta) = \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2}\frac{(\theta-\mu_0)^2}{\sigma_0^2}}$$

By considering only the terms involving $\theta$ we have

$$f_\theta(x) \propto e^{-\frac{1}{2}(x-\theta)^2} \quad \text{and} \quad \pi_0(\theta) \propto e^{-\frac{1}{2}\frac{(\theta-\mu_0)^2}{\sigma_0^2}}$$

This means

$$\pi_1(\theta|x) \propto e^{-\frac{1}{2}(x-\theta)^2} \cdot e^{-\frac{1}{2}\frac{(\theta-\mu_0)^2}{\sigma_0^2}} = \exp\left(-\frac{1}{2}\left((x-\theta)^2 + \frac{(\mu_0-\theta)^2}{\sigma_0^2}\right)\right)$$

Consider just the term of the exponent involving $\theta$

$$\begin{aligned}
& (x-\theta)^2 + \frac{(\mu_0-\theta)^2}{\sigma_0^2} \\
= \; & x^2 - 2x\theta + \theta^2 + \frac{1}{\sigma_0^2}\left(\mu_0^2 - 2\mu_0\theta + \theta^2\right) \\
\propto \; & -2x\theta + \theta^2 + \frac{1}{\sigma_0^2}\left(-2\mu_0\theta + \theta^2\right) \\
= \; & \frac{1}{\sigma_0^2}\left[-2\sigma_0^2 x\theta + \sigma_0^2\theta^2 - 2\mu_0\theta + \theta^2\right] \\
= \; & \frac{1}{\sigma_0^2}\left[\theta^2(1+\sigma_0^2) - 2\theta(\mu_0 + x\sigma_0^2)\right] \\
= \; & \frac{1+\sigma^2}{\sigma_0^2}\left[\theta^2 - 2\theta\left(\frac{\mu_0 + x\sigma_0^2}{1+\sigma_0^2}\right)\right] \\
\propto \; & \frac{1+\sigma^2}{\sigma_0^2}\left(\theta - \left(\frac{\mu_0 + x\sigma_0^2}{1+\sigma_0^2}\right)\right)^2 \text{ by completing the square}
\end{aligned}$$

Substituting this result back into the expression for the posterior gives

$$\begin{aligned}\pi_1(\theta|x) &\propto \exp\left(-\frac{1}{2}\cdot\frac{\left(\theta-\left(\frac{\mu_0+x\sigma_0^2}{1+\sigma_0^2}\right)\right)^2}{\sigma_0^2/(1+\sigma_0^2)}\right)\\ &\sim \text{Normal}\left(\frac{\mu_0+x\sigma_0^2}{1+\sigma_0^2},\frac{\sigma_0^2}{1+\sigma_0^2}\right)\end{aligned}$$

Thus $\pi_1 \sim \text{Normal}(\mu_1,\sigma_1^2)$ where

$$\mu_1 := \frac{\mu_0+x\sigma_0^2}{1+\sigma_0^2} \quad\text{and}\quad \sigma_1^2 := \frac{\sigma_0^2}{1+\sigma_0^2}$$

**Question 5)**
Use the answer to the 4) to formulate a version of the Thompson sampling algorithm for a two-armed bandit, where the rewards from arm $i$ are idd normally distribution with unknown mean $\mu_i$ and unit variance. Explain your algorithm in sufficient detail to enable a non-expert to implement it. You may assume that the non-expert has access to a software package that will generate independent random variables with specified parameters, from any of a family of commonly used probability distributions.

**Answer 5)**
Consider a two-armed bandit where the rewards for each arm are modelled by IID random variables $X_1, X_2$ each with distribution $\text{Normal}(\mu_i, 1)$ with means $\mu_1, \mu_2$ unknown.

Here I give a version of the Thompson Sampling algorithm for solving the multi-armed bandit problem for this bandit, with a round limit $T$.

I. Define a $\text{Normal}(\mu_0,\sigma_0^2)$ prior for the mean of each arm, with the values of $\mu_0,\sigma_0^2$ chosen arbitrarily. (Perhaps $\mu_0=0,\sigma_0^2=1$).

II. To start the $t^{th}$ round, sample $\hat{\mu}_1(t)$ from the prior for arm one and $\hat{\mu}_2(t)$ from the prior for arm two.

III. If $\hat{\mu}_1(t)\geq\hat{\mu}_2(t)$ then play arm one; otherwise, play arm two. Let $x$ denote the observed reward from the played arm.

IV. Suppose the prior for the mean of the played arm at the start of the $t^{th}$ round was $\text{Normal}(\mu_t,\sigma_t^2)$. Define the posterior for the mean of the played arm to be $\text{Normal}(\mu_{t+1},\sigma_{t+1}^2)$ where

$$\mu_{t+1} := \frac{\mu_t+x\sigma_t^2}{1+\sigma_t^2} \quad\text{and}\quad \sigma_{t+1}^2 := \frac{\sigma_t^2}{1+\sigma_t^2}$$

V. For the non-played arm, define the posterior for its mean to be same at its prior at the the start of the $t^{th}$ round.

VI. Repeat steps II.-V. until $T$ rounds have been played, using the posteriors produced in round $t$ as the priors for round $t+1$.