Stochastic Optimisation - Assessed Problem Sheet 1

Dom Hutchinson

October 30, 2020

Question 1.
There are 100 attendees at a party, and they drink 3.5 glasses of champagne on average. Assuming that everybody drinks at least one glass of champagne, find the best upper bound that you can on the number of attendees who drink 6 or more glasses of champagne.

What is the best lower bound you can find? Justify your answer briefly.

Answer 1.
Let $X_1, \dots, X_n$ be IID random variables each modelling how many drinks each attendee drank after their first (We assume all attendees have at least one drink). For this scenario we have that $\mathbb{E}[X_i] = \frac{7}{2} - 1 = \frac{5}{2}$ and $X_i \geq 0$ for all i.

Thus, by *Markov's Inequality*, we have that

$$\forall\ i \in [1, 100] \quad \mathbb{P}(X_i \geq 6 - 1) \leq \frac{\mathbb{E}[X_i]}{6-1} = \frac{5/2}{5} = \frac{1}{2}$$

Then, out of the 100 attendees at most $100 \times \frac{1}{2} = 50$ attendees had at least 6 drinks.

Consider a scenario where fifty guests drink 3 drinks and the other fifty drink 4 drinks. This scenarios fulfils the requirements that all attendees have at least one drink and on average each has $\frac{7}{2}$ drinks. In this scenario no attendees have at least 6 drinks, thus the lower bound for the number of attendees who have at least 6 drink is 0.

The number of attendees who have at least 6 drinks is between 0 and 50.

Question 2.
Consider a Bayesian approach to the problem of inferring the mean of a Poisson distribution with unknown mean. More precisely, suppose $X \sim \text{Poisson}(\lambda)$, where λ is unknown. Let π_0 denote the prior distribution of λ. Let $\pi_1(\cdot|n)$ denote the posterior distribution, conditional on observing a sample of X which takes the value n. In other words

$$\pi_1(\lambda|n) \propto \pi_0(\lambda)p_\lambda(n)$$

where p_λ denotes the probability mass function of a Poisson(λ) random variable.

Show that, if we take π_0 to be a Gamma distribution with shape parameter α and rate parameter β, denoted Gamma(α, β), then $\pi_1(\cdot|n)$ is the density of a Gamma($\alpha + n, \beta + 1$) random variable.

Answer 2.
Let $X \sim \text{Poisson}(\lambda)$ where λ is unknown, π_0 be the prior distribution for λ and $\pi_1(\cdot|n)$ be the posterior distribution for λ, given the value n was sampled from X. This means

$$\pi_1(\lambda|n) \propto \pi_0(\lambda)p_\lambda(n)$$

where $p_\lambda(n) := \mathbb{P}(X = n)$ given $X \sim \text{Poisson}(\lambda)$.
Suppose $\pi_0 \sim \text{Gamma}(\alpha, \beta)$ and note that

$$p_\lambda(n) = \frac{\lambda^n e^{-\lambda}}{n!} \quad \text{and} \quad \pi_0(\lambda) = \frac{\lambda^{\alpha-1}e^{-\lambda\beta}}{\Gamma(\alpha)\beta^\alpha}$$

By consider only the terms involving λ we have that

$$p_\lambda(n) \propto \lambda^n e^{-\lambda} \quad \text{and} \quad \pi_0(\lambda) \propto \lambda^{\alpha-1}e^{-\lambda\beta}$$

Now consider the distribution of the posterior $\pi_1(\cdot|n)$

$$\begin{array}{rcl}\pi_1(\lambda|n) & \propto & \pi_0(\lambda)p_\lambda(n) \\ & \propto & \left(\lambda^{\alpha-1}e^{-\lambda\beta}\right)\cdot\left(\lambda^n e^{-\lambda}\right) \\ & = & \lambda^{n+\alpha-1}e^{-\lambda(\beta+1)}\end{array}$$

By comparing this expression to that of a Gamma distribution we have that

$$\pi_1(\lambda|n) \sim \text{Gamma}(\alpha + n, \beta + 1)$$

Question 3.
Use the answer to the last question to formulate a version of the Thompson sampling algorithm for a two-armed bandit, where the rewards from arm I are IID Poisson(λ_i) random variables, with unknown means λ_1, λ_2. The rewards from the two arms are independent.

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 commonly used probability distributions.

Answer 3.
Consider a two-armed bandit where the rewards from each arm are modelled by IID random variables X_1, X_2 each with distribution Poisson(λ_i) with means λ_1, λ_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 Gamma(α, β) distribution prior for the mean of each arm, with the values of α, β chosen arbitrarily (Perhaps $\alpha = \beta = 1$).

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 n denote the observed reward from the played arm.

IV. Suppose the prior for the mean of the played arm at the start of this round was a Gamma(α_t, β_t) distribution. Define the posterior for the mean of the played arm to be a Gamma($\alpha_t + n, \beta_t + 1$) distribution.

V. For the non-played arm, define the posterior for its mean to be the same as its prior at the start of this round.

VI. Repeat steps II.-V. until T rounds have been played. Use the posteriors from round t as the priors for round $t + 1$.

N.B. After round t the posterior for the mean of arm i will be a Gamma$(\alpha + Y_t, \beta + N_t)$ distribution where Y_t is the sum-total reward received from playing arm i in the first t rounds and N_t is the number of times arm i was played in the first t rounds.