1 / 1
points

**1.**

Prior is said to be conjugate to a likelihood function if:

- the prior, the likelihood function and the posterior would be in a same family of distributions
- the posterior would stay in the same family of distributions as prior

**Correct**
Posterior and prior are both distributions over $\theta$, so they can lie in the same family

- the prior is from the same family of distributions as the likelihood
- the prior lies in the same family of distributions as the likelihood

1 / 1
points

**2.**

Finding a conjugate prior is useful because:

- It leads to a better MAP estimate

**Un-selected is correct**

- We can perform analytical inference and find posterior distribution instead of taking point MAP estimate

**Correct**
Since posterior lies in a known family of distributions, we will be able to perform analytical inference

- It is the only prior for which it is possible to perform analytical inference

**Un-selected is correct**

- As long as posterior will stay in the same family with prior, the integral $p(x_{new} \mid x) = \int p(x_{new} \mid \theta) p(\theta \mid x) d\theta$ which is used for prediction is also tractable

**Correct**
This integral is called the evidence and it can be computed analytically if prior, likelihood and posterior are known

1 / 1
points

**3.**

Out of the following pairs of priors and likelihood functions, choose those that are conjugate:

- [ ] $\Gamma(\lambda \mid \alpha, \beta)$ prior over parameter $\lambda$ of $Exp(x \mid \lambda)$ likelihood ($\Gamma(x, \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ and $Exp(x \mid \lambda) = \lambda e^{-\lambda x}$)

**Correct**
Multiplying these distribution and grouping the terms will lead to gamma distribution again

- [ ] $\mathcal{N}(\mu_1 \mid m, s^2)$ prior over parameter $\mu_1$ for $\mathcal{N}(X \mid \mu_1, \sigma_1^2)$ likelihood

**Correct**
This example was discussed in a lecture

- [ ] $\Gamma(\sigma_1^2 \mid \alpha, \beta)$ prior over parameter $\sigma_1^2$ of $\mathcal{N}(X \mid \mu_1, \sigma_1^2)$ likelihood

**Un-selected is correct**

- [ ] $\mathcal{N}(\sigma_1^2 \mid m, s^2)$ prior over parameter $\sigma_1^2$ of $\mathcal{N}(X \mid \mu_1, \sigma_1^2)$ likelihood

**Un-selected is correct**

---

1 / 1
points

**4.**

---

Which of the following prior distributions over parameter $\sigma^2$ are conjugate to likelihood $\mathcal{N}(x \mid \mu, \sigma^2)$?

- [ ] Inverse gamma with pdf $p(\sigma^2 \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)(\sigma^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right)}$

**Correct**
Multiplying these distribution and grouping the terms will lead to normal distribution

- [ ] $Exp(\sigma^2 \mid \lambda) = \lambda e^{-\lambda \sigma^2}$

**Un-selected is correct**

- [ ] Scaled inverse chi-squared with pdf $f(\sigma^2 \mid \nu, \tau) = \frac{\left(\tau^2 \nu/2\right)^{\nu/2}}{\Gamma(\nu/2)} \frac{\exp\left(-\frac{\nu\tau^2}{2\sigma^2}\right)}{(\sigma^2)^{1+\nu/2}}$

**Correct**
Multiplying these distribution and grouping the terms will lead to normal distribution

- [ ] $\mathcal{N}(\sigma^2 \mid \mu_1, \sigma_1^2)$

**Un-selected is correct**

---

1 / 1
points

**5.**

Choose the correct statements:

- [ ] For arbitrary likelihood and prior pair, we can always perform inference and compute posterior analytically

**Un-selected is correct**

- [ ] Although not for every pair of prior and likelihood there is an analytical expression for posterior, we can always find a conjugate prior in some simple family and compute posterior analytically

**Un-selected is correct**

- [ ] Putting initial knowledge into prior distribution is an advantage of Bayesian approach

**Correct**
That's the one

- [ ] For some problems conjugate prior may be inadequate

**Correct**
That's true

1 / 1
points

**6.**

Imagine that you want to pat your friend's cat Becky. Cats are really random creatures.

Becky might get grumpy and scratch you with probability $p$ or curl up and start purring (with prob. $1 - p$). You don't know Becky well yet, so you estimate prior on $p$ to be distributed as $Beta(2, 2)$. Within one evening, Becky has scratched you 6 times and only 2 times she purred. What will be the parameters for posterior distribution over $p$? What is the MAP-estimate for $p$?

Enter your answers separated by comma: e.g. if you think that correct answer is $Beta(1, 0.2)$ and MAP is 3, you should enter 1,0.2,3. Express real numbers as decimals with dot as delimiter.

Enter answer here

**Correct Response**