

Bayesian Inference: Assignment 2

Answers to the starred (*) questions should be submitted to Canvas by no later than **4pm, Friday 8th March 2024**. Solutions to non-assessed questions will be made available in class and/or in Canvas.

Question 1*

Re-visit question 1 from assignment 1. Recall that we were interested in θ , the probability that particulate matter in the air as a result of a dust storm, results in an environmental health warning being issued.

(a) This question concerns Bayesian confidence intervals:

- (i) Assuming the informative prior used in assignment 1 (not the ignorance prior), which of the following represent valid 95% posterior Bayesian confidence intervals for θ ? Explain your answer. **[3 marks]**

(0.2463, 0.5331) (0.1342, 0.4359) (0.1102, 0.5021) (0.1422, 0.4461)

- (ii) Which interval listed in part (a) is a 95% posterior highest density interval for θ ? Explain your answer. **[2 marks]**
- (iii) Find the 95% prior highest density interval for θ , and compare it to that obtained from the posterior in part (ii). **[2 marks]**

(b) This question is about the Bayesian posterior predictive distribution:

- (i) For the next ten major dust storms to hit Greece and Cyprus this year, find the predictive probability distribution for the number Y that will result in an environmental health warning. **[9 marks]**

Hint: The random variable X follows a Beta-binomial $BetaBin(n, a, b)$ distribution (n positive integer, $a > 0, b > 0$) if it has probability function

$$f(x|n, a, b) = \binom{n}{x} \frac{B(x+a, b+n-x)}{B(a, b)}, \quad x = 0, 1, \dots, n,$$

where $B(a, b)$ is the beta function and

$$B(a, b) = \frac{(a-1)!(b-1)!}{(a+b-1)!}, \quad \text{for } a \text{ and } b \text{ positive integers.}$$

- (ii) Use your answer to part (i) to find a 95% prediction set for the number of dust storms (of the next ten observed) to trigger an environmental health warning. **[4 marks]**

Question 2

The negative binomial distribution is used to model scenarios in which we observe the number of independent (success-fail) trials needed before we see a successful trial; the trials must have the same success probability. Suppose $x_1, x_2, \dots, x_n$ are a random sample from a negative binomial $NegBin(k, \theta)$ distribution, where k is known.

- (i) Verify that the conjugate prior distribution is a Beta distribution.
- (ii) Determine the choice of parameters g and h for the $Beta(g, h)$ distribution that give it maximal variance.
Hint: reparameterise the distribution in terms of its mean $m = g/(g+h)$ and $s = g+h$; determine the choice of m and s that maximises the variance of the distribution, and hence the choice of g and h .

Determine the posterior distribution for θ assuming

- (iii) vague prior knowledge;
- (iv) a very large sample.

Question 3

The Rayleigh distribution is often used to measure variability in magnetic resonance imaging (MRI). Suppose $x_1, x_2, \dots, x_n$ are a random sample from a Rayleigh $R(\theta)$ distribution. Determine the posterior distribution for θ assuming

- (i) vague prior knowledge;
- (ii) a very large sample.

Question 4

There were 4 “significant” flooding events in the UK in 2022. The posterior predictive distribution for Y , the number of significant flooding events in 2023, is shown below:

y	0	1	2	3	4	5	6	7	8	9	10
$f(y x=4)$	0.00	0.01	0.04	0.10	0.13	0.45	0.17	0.05	0.03	0.01	0.01

- (a) Find the most plausible 90% prediction set for Y .
- (b) Find $E(y|x=4)$ and $SD(y|x=4)$.

Solutions

Question 1

(a) Bayesian confidence intervals:

- (i) Recall from question 1 in assignment 1 that the posterior for θ is $\theta|\mathbf{x} \sim \text{Be}(9, 23)$. A valid 95% Bayesian confidence interval (a, b) for $\theta|\mathbf{x}$ would be such that $\Pr(a < \theta < b|\mathbf{x}) = 0.95$. We can easily check this in R for each of the intervals provided:

```
> pbeta(0.5331, 9, 23) - dbeta(0.2463, 9, 23)
[1] 0.6501951
> pbeta(0.4359, 9, 23) - dbeta(0.1342, 9, 23)
[1] 0.9499906
> pbeta(0.5021, 9, 23) - dbeta(0.1102, 9, 23)
[1] 0.9899973
> pbeta(0.4461, 9, 23) - dbeta(0.1422, 9, 23)
[1] 0.9500078
```

Thus, both $(0.1342, 0.4359)$ and $(0.1422, 0.4461)$ are valid 95% Bayesian confidence intervals. **[1 mark for correct identification, 2 marks for reasoning]**

- (ii) We know from the course that a highest density interval also has equal density (γ) at both a and b . Again, we can check in R (or we could do so by hand!). For the first interval:

```
> dbeta(0.1342, 9, 23)
[1] 0.8015702
> dbeta(0.4359, 9, 23)
[1] 0.8010577
```

And for the second interval:

```
> dbeta(0.1422, 9, 23)
[1] 1.038543
> dbeta(0.4461, 9, 23)
[1] 0.6451823
```

Thus, only the first interval $(0.1342, 0.4359)$ satisfies $\pi(\theta = a|\mathbf{x}) = \pi(\theta = b|\mathbf{x}) = \gamma$, and so $(0.1342, 0.4359)$ is the highest density interval. **[2 marks]** We could check this using the `nclbayes` package:

```
> hdiBeta(0.95, 9, 23)
[1] 0.1341826 0.4358943
```

- (iii) We could use the `nclbayes` package:

```
> hdiBeta(0.95, 5, 17)
[1] 0.06921858 0.39949020
```

Thus, the prior HDI is (0.069, 0.399) [1 mark] – similar to that for the posterior (it overlaps!), but shifted to the left – suggesting we are now slightly more inclined to believe that a dust storm will trigger a health warning. [1 mark]

(b) Predictive distribution:

- (i) Suppose the number of dust storms triggering a warning in the future is $Y|\theta \sim \text{Bin}(10, \theta)$. The predictive probability function of Y is, for $y = 0, 1, 2, \dots, 10$,

$$\begin{aligned}
 f(y|x=4) &= \int_{\Theta} f(y|\theta) \pi(\theta|x=4) d\theta \\
 &= \int_0^1 \binom{10}{y} \theta^y (1-\theta)^{10-y} \times \frac{\theta^8 (1-\theta)^{22}}{B(9, 23)} d\theta \\
 &= \binom{10}{y} \frac{1}{B(9, 23)} \int_0^1 \theta^{y+8} (1-\theta)^{32-y} d\theta \quad [2 \text{ marks}] \\
 &= \binom{10}{y} \frac{B(y+9, 33-y)}{B(9, 23)} \int_0^1 \frac{\theta^{(y+9)-1} (1-\theta)^{(33-y)-1}}{B(y+9, 33-y)} d\theta \quad [4 \text{ marks}] \\
 &= \binom{10}{y} \frac{B(y+9, 33-y)}{B(9, 23)}. \quad [1 \text{ mark}]
 \end{aligned}$$

That is, $Y|x=4 \sim \text{BetaBin}(10, 9, 23)$. [2 mark]

- (ii) Probably best to write an R function?

```

> betabin = function(y, n, a, b){
  numerator = beta(y + a, n + b - y)
  denominator = beta(a, b)
  answer = choose(10,y)*numerator/denominator
  return(answer)}

> cbind(0:10,round(betabin(0:10, 10, 9, 23),3))
  [,1] [,2]
[1,]  0 0.058
[2,]  1 0.162
[3,]  2 0.235
[4,]  3 0.230
[5,]  4 0.166
[6,]  5 0.093
[7,]  6 0.040
[8,]  7 0.013
[9,]  8 0.003
[10,] 9 0.000
[11,] 10 0.000

```

The above is our predictive distribution, that is, $f(y|x=4)$. For the prediction set, we need to include values for y with greatest predictive probability, such that the

sum of their probabilities is close to 0.95. We include, in order (and with cumulative probabilities in parentheses): $[(2(0.235), 3(0.465), 4(0.631), 1(0.793), 5(0.886), 0(0.944)]$, i.e. $[0, 5]$.

[3 marks for method, 1 mark for final answer of $[0, 5]$.

Question 2

- (i) If the prior distribution is $\theta \sim \text{Beta}(g, h)$ then, using Bayes Theorem, the posterior density is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \theta^{g-1}(1-\theta)^{h-1} \times \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i}, \quad 0 < \theta < 1 \\ &\propto \theta^{g+n\bar{x}-1}(1-\theta)^{h+n-n\bar{x}-1}, \quad 0 < \theta < 1 \\ &\text{i.e. } \theta|\mathbf{x} \sim \text{Beta}(G = g + nk, H = h + n\bar{x} - nk).\end{aligned}$$

Therefore, both the prior and posterior distributions are Beta distributions and so the Beta is the conjugate prior distribution.

- (ii) As the beta distribution restricts values to the range $(0,1)$, there is a finite upper limit to the variance. If $\theta \sim \text{Beta}(g, h)$ then it has variance $\text{Var}(\theta) = gh/\{(g+h)^2(g+h+1)\}$. Note that this distribution requires its parameters $g > 0$ and $h > 0$. If we now transform the parameters to $m = g/(g+h)$ and $s = g+h$ then $\text{Var}(\theta) = m(1-m)/(s+1)$. Note that, as this is a 1-1 transformation, we have a free choice of values for m and s subject to the restrictions $0 < m < 1$ and $s > 0$. Clearly, this ratio is maximised when the numerator is maximised and the denominator is minimised. Now $m(1-m)$ is maximised at $m = 1/2$, and $(s+1)$ is minimised by taking s as small as possible. In terms of g and h , these conditions are $g = h$ and $g+h \rightarrow 0$, that is, $g = h \rightarrow 0$.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g = h \rightarrow 0$.

- (iii) We make the Beta prior distribution vague by making it have as large a variance as possible. This is achieved, from (ii) above, by letting $g = h \rightarrow 0$. Hence, using the result in (i), the posterior distribution assuming vague prior information is

$$\theta|\mathbf{x} \sim \text{Beta}(nk, n\bar{x} - nk).$$

- (iv) The asymptotic posterior distribution (as $n \rightarrow \infty$) is

$$\theta|\mathbf{x} \sim N(\hat{\theta}, J(\hat{\theta})^{-1}),$$

where

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = \frac{nk}{\theta^2} + \frac{n(\bar{x} - k)}{(1-\theta)^2}.$$

Now

$$\begin{aligned}
 \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 &\implies \frac{nk}{\hat{\theta}} - \frac{n(\bar{x} - k)}{1 - \hat{\theta}} = 0 \\
 &\implies \hat{\theta} = \frac{k}{\bar{x}} \\
 &\implies J(\hat{\theta}) = \frac{n\bar{x}^3}{k(\bar{x} - k)} \\
 &\implies J(\hat{\theta})^{-1} = \frac{k(\bar{x} - k)}{n\bar{x}^3}.
 \end{aligned}$$

Therefore, for large n , the posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(\frac{k}{\bar{x}}, \frac{k(\bar{x} - k)}{n\bar{x}^3}\right) \quad \text{approximately.}$$

Question 3

(i) The likelihood function is

$$f(\mathbf{x}|\theta) = \prod_{i=1}^n 2x_i\theta e^{-\theta x_i^2} \propto \theta^n \exp\left\{-\theta \sum_{i=1}^n x_i^2\right\} = \theta^n \exp\{-n\bar{x}^2\theta\},$$

where $\bar{x}^2 = \sum_{i=1}^n x_i^2/n$ is the mean of the squared observations. Therefore the conjugate prior distribution is a Gamma $Ga(g, h)$ distribution. Using Bayes Theorem, the posterior density is

$$\begin{aligned}
 \pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\
 &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \theta^n e^{-n\bar{x}^2\theta}, \quad \theta > 0,
 \end{aligned}$$

that is, $\theta|\mathbf{x} \sim Ga(G = g + n, H = h + n\bar{x}^2)$.

We saw in Example 1.8 that to make a gamma distribution vague, we take $g \rightarrow 0$ and $h \rightarrow 0$. Therefore using a vague prior distribution gives the posterior distribution as $\theta|\mathbf{x} \sim Ga(n, n\bar{x}^2)$.

(ii) The asymptotic posterior distribution (as $n \rightarrow \infty$) is

$$\theta|\mathbf{x} \sim N(\hat{\theta}, J(\hat{\theta})^{-1}),$$

where

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = \frac{n}{\theta^2}.$$

Now

$$\begin{aligned}
 \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 &\implies \frac{n}{\hat{\theta}} - n\bar{x}^2 = 0 \\
 &\implies \hat{\theta} = 1/\bar{x}^2 \\
 &\implies J(\hat{\theta})^{-1} = \frac{1}{n} (\bar{x}^2)^2.
 \end{aligned}$$

Therefore, for large n , the posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(1/\bar{x^2}, \frac{1}{n}(\bar{x^2})^2\right) \quad \text{approximately.}$$

Note that the posterior distribution is centred on the likelihood mode and its variance tends to zero as $n \rightarrow \infty$.

Question 4

(a) We have:

y	Cumulative probability	Decision?
5	0.45	Keep going
6	$0.45 + 0.17 = 0.62$	Keep going
4	$0.62 + 0.13 = 0.75$	Keep going
3	$0.75 + 0.10 = 0.85$	Keep going
7	$0.85 + 0.05 = 0.90$	Stop

So the 90% prediction set is $\{5, 6, 4, 3, 7\}$, in other words $\{3, 7\}$.

(b) We have

$$E(y|x=4) = 1 \times 0.01 + 2 \times 0.04 + \dots + 10 \times 0.01 = 4.96.$$

We also have

$$E(y^2|x=4) = 1^2 \times 0.01 + 2^2 \times 0.04 + \dots + 10^2 \times 0.01 = 26.7.$$

Therefore

$$\text{Var}(y|x=4) = 26.7 - 4.96^2 = 2.0984,$$

$$\text{giving } SD(y|x=4) = \sqrt{2.0984} = 1.4486.$$

Therefore,

$$\begin{aligned} f(Y|x=4) &= \binom{10}{y} \frac{B(y+9, 33-y)}{B(9, 23)} \\ &= \binom{10}{y} \frac{\frac{(y+9-1)!(33-y-1)!}{(y+9+33-y-1)!}}{\frac{(2-1)!(23-1)!}{(2+23-1)!}} \\ &= \binom{5}{y} \times \frac{(y+1)!(27-y)! \times 24!}{22! \times 29!}. \end{aligned}$$

So, for example,

$$P(Y=0|x=1) = \binom{5}{0} \times \frac{1! \times 27! \times 24!}{22! \times 29!} = 0.680 \quad (\text{See Table 4.4})$$

Distributions for data

Binomial distribution

If $X|\theta \sim \text{Bin}(k, \theta)$ then it has probability function

$$f(x|\theta) = \binom{k}{x} \theta^x (1 - \theta)^{k-x}, \quad x = 0, 1, \dots, k,$$

where k is a positive integer and $0 < \theta < 1$. Also, $E(X) = k\theta$ and $\text{Var}(X) = k\theta(1 - \theta)$.

Exponential distribution

If $X|\lambda \sim \text{Exp}(\lambda)$ then it has density

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x > 0$$

where $\lambda > 0$. Also, $E(X) = 1/\lambda$ and $\text{Var}(X) = 1/\lambda^2$.

Gamma distribution

If $X|\alpha, \lambda \sim \text{Ga}(\alpha, \lambda)$ then it has density

$$f(x|\alpha, \lambda) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0,$$

where $\alpha > 0$ and $\lambda > 0$. Also, $E(X) = \alpha/\lambda$ and $\text{Var}(X) = \alpha/\lambda^2$.

Inverse Gaussian distribution

If $X|\mu, \lambda \sim \text{IG}(\mu, \lambda)$ then it has density

$$f(x|\mu, \lambda) = \frac{\sqrt{\lambda}}{\sqrt{2\pi x^3}} \exp \left\{ -\frac{\lambda(x - \mu)^2}{2\mu^2 x} \right\}, \quad x > 0,$$

where $\mu > 0$ and $\lambda > 0$. Also, $E(X) = \mu$ and $\text{Var}(X) = \mu^3/\lambda$.

Laplace distribution

If $X|\mu, \sigma \sim \text{La}(\mu, \sigma^2)$ then it has density

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2}\sigma} \exp \left\{ -\frac{\sqrt{2}|x - \mu|}{\sigma} \right\}, \quad x > 0,$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Also, $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$.

Log-normal distribution

If $X|\mu, \sigma \sim LN(\mu, \sigma^2)$ then it has density

$$f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\log x - \mu)^2\right\}, \quad x > 0$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Also, $E(X) = e^{\mu+\sigma^2/2}$, $Var(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$. Further $E(\log X) = \mu$ and $Var(\log X) = \sigma^2$.

Negative binomial distribution

If $X|\theta \sim NegBin(k, \theta)$ then it has probability function

$$f(x|\theta) = \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k}, \quad x = k, k+1, \dots,$$

where k is a positive integer and $0 < \theta < 1$. Also $E(X) = k/\theta$ and $Var(X) = k(1-\theta)/\theta^2$.

Normal distribution

If $X|\mu, \tau \sim N(\mu, 1/\tau)$ then it has density

$$f(x|\mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x-\mu)^2\right\}, \quad x \in \mathbb{R}$$

where $\mu \in \mathbb{R}$ and $\tau > 0$. Also, $E(X) = \mu$ and $Var(X) = 1/\tau$. The distribution has the following quantiles

x	1.2816	1.6449	1.9600	2.3263	2.5758
$\Pr(X < x)$	0.9	0.95	0.975	0.99	0.995

Pareto distribution

If $X|\theta, \alpha \sim Pa(\theta, \alpha)$ then it has density

$$f(x|\alpha, \theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \quad x > \theta$$

where $\alpha > 0$ and $\theta > 0$. Also $E(X) = \alpha\theta/(\alpha-1)$, $\alpha > 1$ and $Var(X) = \alpha\theta^2/\{(\alpha-1)^2(\alpha-2)\}$, $\alpha > 2$.

Poisson distribution

If $X|\theta \sim Po(\theta)$ then it has probability function

$$f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, \dots,$$

where $\theta > 0$. Also, $E(X) = \theta$ and $Var(X) = \theta$.

Rayleigh distribution

If $X|\theta \sim R(\theta)$ then it has density

$$f(x|\theta) = 2x\theta e^{-\theta x^2}, \quad x > 0,$$

where $\theta > 0$. Also $E(X) = \sqrt{\pi/(4\theta)}$ and $Var(X) = (4 - \pi)/(4\theta)$.

Uniform distribution

If $X|\phi, \theta \sim U(\phi, \theta)$ then it has density

$$f(x|\phi, \theta) = \frac{1}{\theta - \phi}, \quad \phi < x < \theta,$$

where $\phi < \theta$. Also, $E(X) = (\phi + \theta)/2$ and $Var(X) = (\theta - \phi)^2/12$.

von Mises distribution

If $X|\mu, \lambda \sim vM(\mu, \lambda)$ then it has density

$$f(x|\mu, \lambda) = \frac{1}{2\pi I_0(\lambda)} \exp\{\lambda \cos(x - \mu)\}, \quad 0 \leq x < 2\pi$$

where $0 \leq \mu < 2\pi$ and $\lambda > 0$. Also $I_0(\cdot)$ is the modified Bessel function of order zero.

Weibull distribution

If $X|\beta, \lambda \sim Wei(\beta, \lambda)$ then it has density

$$f(x|\beta, \lambda) = \beta \lambda^\beta x^{\beta-1} e^{-\lambda x^\beta}, \quad x > 0,$$

where $\beta > 0$ and $\lambda > 0$. Also, $E(X) = \Gamma(1 + 1/\beta)/\lambda$ and $Var(X) = \Gamma(1 + 2/\beta)/\lambda^2 - E(X)^2$.

Distributions for prior beliefs

Beta distribution

If $\theta \sim \text{Beta}(g, h)$ then it has density

$$\pi(\theta) = \frac{\theta^{g-1}(1-\theta)^{h-1}}{B(g, h)}, \quad 0 < \theta < 1,$$

where $g > 0$ and $h > 0$. Also, $E(\theta) = g/(g+h)$ and $\text{Var}(\theta) = gh/\{(g+h)^2(g+h+1)\}$.

Exponential distribution

If $\theta \sim \text{Exp}(h)$ then it has density

$$\pi(\theta) = he^{-h\theta}, \quad \theta > 0,$$

where $h > 0$. Also, $E(\theta) = 1/h$ and $\text{Var}(\theta) = 1/h^2$.

Gamma distribution

If $\theta \sim \text{Ga}(g, h)$ then it has density

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0,$$

where $g > 0$ and $h > 0$. Also, $E(\theta) = g/h$ and $\text{Var}(\theta) = g/h^2$.

Generalised t distribution

If $\mu \sim t_a(b, c)$ then it has density

$$\pi(\mu) = \frac{\Gamma(\frac{a+1}{2})}{\sqrt{ac}\pi\Gamma(\frac{a}{2})} \left\{ 1 + \frac{(\mu-b)^2}{ac} \right\}^{-\frac{a+1}{2}}, \quad \mu \in \mathbb{R},$$

where $b \in \mathbb{R}$, $a > 0$ and $c > 0$. Also, $E(\mu) = b$ and $\text{Var}(\mu) = ac/(a-2)$ if $a \geq 2$.

Inverse Chi distribution

If $\sigma \sim \text{Inv-Chi}(a, b)$ then it has density

$$\pi(\sigma|a, b) = \frac{2b^a \sigma^{-2a-1} e^{-b/\sigma^2}}{\Gamma(a)}, \quad \sigma > 0,$$

where $a > 0$, $b > 0$ and $\Gamma(a)$ is the gamma function. Also $E(\sigma) = \sqrt{b}\Gamma(a-1/2)/\Gamma(a)$ and $\text{Var}(\sigma) = b/(a-1) - E(\sigma)^2$ if $a > 1$. The name of the distribution comes from the fact that $1/\sigma^2 \sim \text{Ga}(a, b) \equiv \chi_{2a}^2/(2b)$.

Log-normal distribution

If $\theta \sim LN(b, c^2)$ then it has density

$$\pi(\theta) = \frac{1}{\sqrt{2\pi} c \theta} \exp \left\{ -\frac{1}{2c^2} (\log \theta - b)^2 \right\}, \quad \theta > 0$$

where $b \in \mathbb{R}$ and $c > 0$. Also, $E(\theta) = e^{b+c^2/2}$, $Var(\theta) = (e^{c^2} - 1)e^{2b+c^2}$. Further $\log \theta \sim N(b, c^2)$ and so $E(\log \theta) = b$ and $Var(\log \theta) = c^2$.

Normal distribution

If $\mu \sim N(b, 1/d)$ then it has density

$$\pi(\mu) = \left(\frac{d}{2\pi} \right)^{1/2} \exp \left\{ -\frac{d}{2} (\mu - b)^2 \right\}, \quad \mu \in \mathbb{R},$$

where $b \in \mathbb{R}$ and $c > 0$. Also, $E(\mu) = b$ and $Var(\mu) = 1/d$.

Normal-gamma distribution

If $\begin{pmatrix} \mu \\ \tau \end{pmatrix} \sim NGa(b, c, g, h)$ then it has density

$$\pi(\mu, \tau) \propto \tau^{g-\frac{1}{2}} \exp \left\{ -\frac{\tau}{2} [c(\mu - b)^2 + 2h] \right\}, \quad \mu \in \mathbb{R}, \tau > 0$$

where $b \in \mathbb{R}$ and $c, g, h > 0$. Also, $\mu|\tau \sim N\left(b, \frac{1}{c\tau}\right)$, $\tau \sim Ga(g, h)$ and has marginal distribution $\mu \sim t_{2g}\left(b, \frac{h}{gc}\right)$.

Uniform distribution

If $\theta \sim U(a, b)$ then it has density

$$\pi(\theta) = \frac{1}{b-a}, \quad a < \theta < b,$$

where $a < b$. Also, $E(\theta) = (a+b)/2$ and $Var(\theta) = (b-a)^2/12$.