We're rewarding the question askers & reputations are being recalculated! Read more



Jeffreys Prior for normal distribution with unknown mean and variance

Asked 4 years, 5 months ago Active 1 year, 7 months ago Viewed 10k times



I am reading up on prior distributions and I calculated Jeffreys prior for a sample of normally distributed random variables with unknown mean and unknown variance. According to my calculations, the following holds for Jeffreys prior:





$$p(\mu,\sigma^2) = \sqrt{det(I)} = \sqrt{det \left(egin{array}{cc} 1/\sigma^2 & 0 \ 0 & 1/(2\sigma^4) \end{array}
ight)} = \sqrt{rac{1}{2\sigma^6}} \propto rac{1}{\sigma^3}.$$



Here, I is Fisher's information matrix.

However, I have also read publications and documents which state

- $p(\mu, \sigma^2) \propto 1/\sigma^2$ see Section 2.2 in Kass and Wassermann (1996).
- $p(\mu, \sigma^2) \propto 1/\sigma^4$ see page 25 in <u>Yang and Berger (1998)</u>

as Jeffreys prior for the case of a normal distribution with unkown mean and variance. What is the 'actual' Jeffreys prior?

bayesian normal-distribution prior jeffreys-prior

edited Jun 11 '15 at 1:15

A. Donda 2,388 10 asked .lun 9 '15 at 18:4



3 Answers

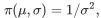


I think the discrepancy is explained by whether the authors consider the density over σ or the density over σ^2 . Supporting this interpretation, the exact thing that Kass and Wassermann write is

6

while Yang and Berger write





$$\pi(\mu,\sigma)=1/\sigma^2,$$

 $\pi(\mu, \sigma^2) = 1/\sigma^4$.

answered Jun 11 '15 at 0:47



- 2 A Thanks, I overlooked this. However, this still does not explain the discrepancy between $1/\sigma^3$ and $1/\sigma^4$. Nussig Jun 11 '15 at 1:02
- Actually, having a prior of $\pi(\mu, \sigma) = 1/\sigma^2$ is the same as having a prior $\pi(\mu, \sigma^2) = 1/\sigma^3$, due the reparametrization property of Jeffreys prior:

$$\pi(\mu,\sigma)=\pi(\mu,\sigma^2)det(J_f)\proptorac{1}{\sigma^3}2\sigma\proptorac{1}{\sigma^2}$$

with J_f the Jacobian matrix of $f:(\mu,\sigma)\to(\mu,\sigma^2)$, i.e.

$$J_f = \left(egin{array}{cc} 1 & 0 \ 0 & 2\sigma \end{array}
ight)$$

- . Nussig Jun 11 '15 at 2:00 /
- 3 riangle @Nussig, I checked the calculation, and I think you are right arriving at $1/\sigma^3$. You are also right that the reparametrization amounts only to a factor $1/\sigma$. Considering this, your calculation is in accordance with Kass and Wassermann, and I can only guess that Yang and Berger made a mistake. This makes sense also since the former is a regular reviewed journal paper and the latter is a draft of a kind of formula collection. - A. Donda Jun 11 '15 at 2:20
- 3 A Kass and Wassermann also note that Jeffreys introduced a modified rule, according to which location and scale parameters should be treated separately. This leads to $\pi(\mu,\sigma)=1/\sigma$ and therefore $\pi(\mu,\sigma^2)=1/\sigma^2$, but still not to $\pi(\mu,\sigma^2)=1/\sigma^4$. – A. Donda Jun 11 '15 at 2:22 \nearrow
- Jim Berger is still an active scientist, so to be sure you might check directly with him: stat.duke.edu/~berger A. Donda Jun 11 '15 at 2:28



The existing answers already well answer the original question. As a physicist, I would just like to add to this discussion a dimensionality argument. If you consider μ and σ^2 to describe a distribution of a random variable in a real 1D space and measured in meters, they have the dimensions $[\mu]\sim m$ and $[\sigma^2] \sim m^2$. To have a physically correct prior, you need it to have the right dimensions, i.e. the only powers of σ physically possible in a **non-parametric** prior are:



$$\pi(\mu,\sigma) \sim 1/\sigma^2$$

and

$$\pi(\mu,\sigma^2)\sim 1/\sigma^3$$

answered Mar 23 '18 at 10:13





Mhy is there σ^3 in the second expression? – cerebrou Feb 25 at 15:59 \nearrow



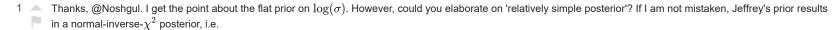
 $\frac{1}{\sigma^3}$ is the Jeffreys prior. However in practice $\frac{1}{\sigma^2}$ is quite often used cause it leads to a relatively simple posterior, the "intuition" of this prior is that it corresponds with a flat prior on $\log(\sigma)$.





answered Jun 10 '15 at 19:41





$$(\mu,\sigma^2)|D\sim \mathcal{N}\chi^{-1}\left(\overline{X},n,n,rac{1}{n}\sum (X_i-\overline{X})^2
ight).$$

The prior $1/\sigma^2$ should result in a normal-inverse- χ^2 posterior, too, just with different parameters. – Nussig Jun 10 '15 at 19:57 \nearrow

Ooh, yes it leads to a normal-inverse- $\chi^2(\bar{X},n,n-1,s^2)$. I just find it more natural that the marginal of σ^2 is an inverse χ^2 with n-1 instead of n degrees of freedom. Anyhow, I certainly did not want to imply that the other priors would lead to annoying distributions. To be honest I didn't know the posterior of the Jeffry's prior by heart nor did I really think to much about it when I wrote the post. - Jorne Biccler Jun 11 '15 at 17:26