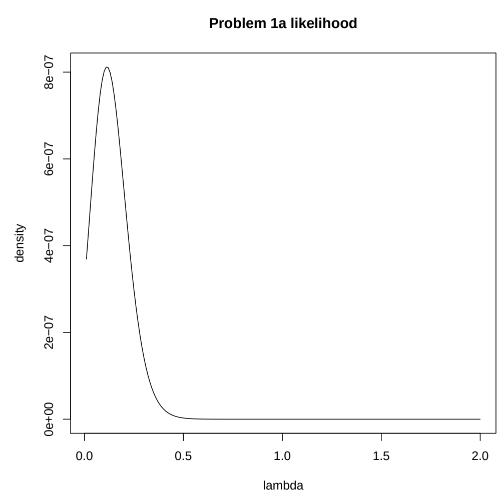
EXERCISES ON BAYESIAN BASICS

Both of the problems below have sections that require use of the computer, to construct plots and to carry out numerical integration. The plot commands in Matlab, Octave or R are straightforward to use. Numerical integration in one dimension (all that is required here) is easy, and you might plow ahead to program it yourself, but it will be easier and more accurate to use the commands that do it automatically. In Octave (and probably also Matlab) there are three univariate integration commands available: `quad`, `quadl`, and `trapz`. They take different approaches to trying to ensure high accuracy, with `quadl` probably the most accurate. In R (and probably S) the function `integrate` is available to do the same thing.

(1) (MacKay example from class) We suppose we have a random sample of individuals who have entered unemployment exactly two weeks before their selection for the sample, the sample has been followed for 20 weeks, and all individuals who remained unemployed for the full 20 weeks have been eliminated from the sample. For each individual i we have data on the completed spell length x_i. We assume spell length follows a gamma distribution, i.e. has pdf $\lambda e^{-\lambda x_i}$.

(a) Write down the likelihood function when the sample is $\{3, 7, 2.5, 19, 11\}$. Plot it.
The likelihood function:

$$\frac{\lambda^5 e^{-42.5\lambda}}{(e^{-2\lambda} - e^{-22\lambda})^5}$$



(b) Show that the claim in lecture that the likelihood function goes to infinity as $\lambda \to 0$ is incorrect.
Both numerator and denominator of the likelihood go to zero as $\lambda \to 0$. However we can apply l'Hôpital's rule (i.e., take derivatives of the numerator and denominator and look at how they behave as $\lambda \to 0$. You would have to apply it five times here, though to reach the answer. A quicker approach is to recognize that the first order expansion of

Date: October 29, 2008.

the expression in parenthesis in the denominator is $-2\lambda + 22\lambda$, so that the lead term in its fifth power is $(20\lambda)^5$. The λ^5 cancels with the λ^5 in the numerator, so that the limit of the ratio is just $1/20^5$, or 3.125e-7. This agrees with the plot.

Some students answered this by just saying the answer is clear from the plot. This is not a good answer. It assumes that between the first point above zero you plotted and 0, the function is nearly a straight line — it is (at least with the 200 points between 0 and 1 that I plotted) but you can't know that without an argument.

(c) Show that the likelihood is integrable.

Both numerator and denominator are continuous, and we know that the ratio is bounded in the neighborhood of zero, by the argument above. The denominator is strictly positive for all positive λ. Therefore the ratio is continuous on $[0, a]$ for any $a > 0$ and therefore bounded on such an interval. A bounded continuous function on a finite interval has a finite integral. What about the integral over $[a, \infty)$ then? The denominator is the fifth power of $e^{-2\lambda}(1 - e^{-20\lambda})$ which exceeds $0.99e^{-2\lambda}$ for $\lambda > 1$. Therefore the likelihood, on the interval $(1, \infty)$, is bounded above by $\lambda^5 e^{-32.5\lambda} .99^{-5}$. This is proportional to a $\mathrm{Gamma}(6, 32.5)$ density, which we know is integrable, so this likelihood, being positive and bounded above by an integrable function, is integrable.

(d) Show that, if we use a flat prior (i.e. treat the normalized likelihood as our pdf for λ) the expected spell length, taking account of parameter uncertainty, is infinite. (For an individual spell with pdf $\lambda e^{-\lambda x}$, conditional on λ, the expected spell length is $1/\lambda$.)

By the law of iterated expectations, the expected spell length is just the expectation over λ of the conditional expected spell lengths given λ. So checking finiteness of this expectation amounts to checking the integrability of the likelihood times $1/\lambda$. This just changes the bounding function for the upper tail that we constructed above from a $\mathrm{Gamma}(6, 32.5)$ to a $\mathrm{Gamma}(5, 32.5)$, so the right tail remains integrable. But in the neighborhood of zero the likelihood itself converges to a non-zero value, so when we multiply it by $1/\lambda$, we get a function that behaves in the neighborhood of zero like $1/\lambda$, which is not integrable. (It has an indefinite integral, $\log \lambda$, so that the integral over the interval (ε, a) is $\log(a) - \log(\varepsilon)$, but since $\log(\varepsilon) \to -\infty$ as $\lambda \to 0$, the integral over $(0, a)$ is $+\infty$.)

Here again some students tried to use graphs or computational results to get an answer, which can't be done, at least by the numerical arguments people were putting forth. Many numerical integration routines give up on any function which is infinite anywhere in the range of integration, yet many such functions are integrable, for example.

(e) Using numerical integration, find the expected spell length if we have a $\mathrm{Gamma}(2, \beta)$ prior pdf on λ — i.e. a prior pdf of $.5\beta^2 \lambda e^{-\beta\lambda}$, with β set at .001, .01, and .1. Is this expected value strongly affected by priors in this range? Are these values of β reasonable choices for a prior on λ in this application to unemployment spells?

The problem gives the wrong form for the $\mathrm{Gamma}(2, \beta)$ pdf: the normalizing constant is $\Gamma(2) = 1! = 1$, not 2!, so the .5 in front was a mistake. The pdf should have been given as simply $\beta^2 \lambda e^{-\beta\lambda}$. Below is R code that I used to do this, all with the mistaken .5 factor in place. However, since we are considering only posterior distributions of a parameter, the scale of the posterior density drops out of the calculation. As we will see later, the scale would matter for model comparison, but we're not doing that here.

The R code defines one function that is the kernel of the posterior, then another that is the integrand for the expectation. The posterior kernel has to be integrated first to find the normalizing constant. It turns out that in its raw form, because numerator and denominator both go very close to zero, the kernel chokes R's `integrate()` function, so I factored e^{-10x} out of numerator and denominator of the kernel, after which `integrate()` works.

The code uses not only the problem's three given values of β .001, .01, and .1, but also 1, 10, 20 and 50. The prior implies zero density at $\lambda = 0$, which is reasonble. We don't believe infinite spell lengths are likely, and this gives us finite posterior expected spell lengths. However the three β's given in the problem have modal λ's of 1000, 100, and 10, respectively, so they all put highest prior likelihood on expected spell lengths of less than one week, even less than one day. Inference is insensitive to priors in this range, but this is only because all three priors imply very strong belief that most spells end before our two week initial cutoff. As can be seen from the results below, if we had used β's in the set $\{1, 10, 20, 50\}$, which might more realistically reflect prior beliefs, we would have found posterior mean spell lengths strongly affected by the prior.

```
> pstrk <- function(x) .5 * bet^2 * x^6 * exp(-(32.5+bet) * x +14.62007)/(1-exp(-20*x))^5
> ek <- function(x) pstrk(x)/x
> bet <- .001
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 7.039455
> bet <- .01
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 7.042629
> bet <- .1
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 7.074433
> bet <- 1
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 7.398644
> bet <- 10
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 11.28445
> bet <- 20
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 16.95790
> bet <- 50
> integrate(ek, 1e-6, Inf)$value/integrate(pstrk, 1e-6, Inf)$value
[1] 40.00224
>
>
```

(f) Explain how it can be that Mackay, parameterizing the exponential as $e^{-x/\alpha}/\alpha$, finds that with a flat prior on α, the likelihood is not integrable, while we, using $\lambda = 1/\alpha$ as the parameter, find it is integrable. If we used a proper prior, could such a reparameterization make the posterior pdf non-integrable? [Hint: These questions are related to the formula for converting a pdf $p(x)$ for x into a pdf $q(y)$ for $y = f(x)$, which involves the Jacobian of f.]

We find that the likelihood converges to a non-zero constant as $\lambda \to 0$. This means that if we used $\alpha = 1/\lambda$ as the parameter, we would have likelihood converging to this same non-zero constant as $\alpha \to \infty$. Thus the likelihood would be non-integrable in α, even though it is integrable in λ. But if we treat $\ell(\lambda)$, the likelihood in λ scaled to integrate to one, as a density (which is what we do when we use a "flat prior" on λ), then translating that density into a density for $\alpha = 1/\lambda$ we have to take account of the Jacobian term, so the density for α becomes $\ell(1/\alpha)\alpha^{-2}$. (To remember this rule, think of a density $f(x)$ always as accompanied by a "dx". So $\ell(\lambda)d\lambda$ becomes $\ell(1/\alpha)\left|\frac{d\lambda}{d\alpha}\right| d\alpha = \ell(1/\alpha)\alpha^{-2}$.) This density for α is integrable. If this Jacobian rule for transforming pdf's to correspond to transformed random variables is followed, the intergrals of densities that are one to start with always stay equal to one under the transformation.

A flat prior on λ, then, corresponds to an improper prior density of $1/\alpha^2$ on α.

(2) A textbook simple model of asset pricing implies that when there is a single asset being saved the price of it in terms of current consumption goods is

$$P_t = E_t\left[\frac{e^r U'(C_{t+1})}{U'(C_t)}\right] ,$$

where $U'()$ is the marginal utility of consumption, E_t is expectation conditional on information at t, and r is the realized yield on the asset, which may be uncertain at date t. If income from savings is the only source of funds at $t+1$, then $C_{t+1} = S_t e^r$, where S_t is savings at t. Economic and financial modelers very commonly assume that U has the "constant relative risk aversion" (CRRA) form $U(C) = C^{1-\gamma}/(1-\gamma)$, because it is convenient and has some intuitively appealing properties.

(a) Suppose $r \sim N(\bar{r}, \sigma^2)$ and that $\bar{r}$ and σ^2 are known. Using the fact that if $x \sim N(\mu, v^2)$ (so that e^x is log-normal), $E[e^x] = e^{\mu + \frac{1}{2}\sigma^2}$, determine how the asset price behaves as the rate of relative risk aversion (γ) and the return variance σ^2 vary.

As a student noted in class, this problem was incompletely specified. I told you to treat S_t as fixed. This corresponds to a simple general equilibrium model in which the assets are in fixed supply ("trees", e.g.) and there is a fixed quantity of consumption goods in the first period. Prices then adjust to the point where individuals, who see themselves as able to trade consumption goods for investment goods, choose to hold exactly the fixed supply S_t of assets. An opposite extreme would be to assume the asset is in perfectly elastic supply, but this would require fixing the supply price P_t, and then only the amount S_t of the asset purchased would vary with changes in the distribution of r. Of course there could be intermediate cases, where the investment good can be produced at increasing cost from the consumption good, but the problem did not give enough information for a solution except under the fixed-S assumption.

The asset price will be

$$\left(\frac{C_t}{S_t}\right)^\gamma E_t[e^{(1-\gamma)r}] = \left(\frac{C_t}{S_t}\right)^\gamma e^{(1-\gamma)\bar{r} + \frac{1}{2}(1-\gamma)^2\sigma^2)}$$

As the return variance σ^2 increases, the price of the asset increases, unless $\gamma = 1$. That this is true when $\gamma < 1$ and risk aversion is therefore low is unsurprising, since the increased variance in r increases the expected return Ee^r. When $\gamma > 1$, the increased value of the asset arises because there is a greater possibility of low consumption in the second period, and having more of the asset when the returns are very low is valuable. (This kind of result depends on there being no risk-free asset available.) What happens as γ varies depends on all the other parameters. The general expression for the derivative of $\log P_t$ with respect to γ is

$$\log(C_t/S_t) - \bar{r} - (1-\gamma)\sigma^2$$

Thus we can be sure that increasing γ decreases the price if $C/S < 1$ and $\gamma < 1$, but otherwise the sign is ambiguous.

(b) MacKay explains in Chapter 23.2 and Chapter 24 how, when $\bar{r}$ and σ^2 are uncertain, inference about them with a conjugate prior (he defines that term) leads to what is called a normal-inverse-gamma posterior joint posterior distribution for $\bar{r}$ and σ^2. Show that with such a distribution for $\bar{r}$ and σ^2 the price of the asset with CRRA utility is infinite. (It may be easiest to show that it is infinite when you integrate over σ^2, for each given $\bar{r}$, which implies that it is infinite when integrated over $\bar{r}$ as well.)

(*This result was noted some years ago by John Geweke. Recently Martin Weitzman has written papers arguing that it can explain most of the "asset pricing paradoxes" that have turned up in empirical finance.*)

The joint distribution for $\bar{r}$ and σ^2 will have a pdf proportional to

$$\sigma^{-2(m+1)} \exp\left(-\frac{(\bar{r}-\hat{r})^2}{2\sigma^2}\right) .$$

For checking finiteness, we are concerned only with the $E[e^{(1-\gamma)r}]$ part of the pricing formula so the integral we would have to evaluate to take the expectation is

$$\int_0^\infty \int_{-\infty}^\infty \sigma^{-2(m+1)} \exp\left(-\frac{(\bar{r}-\hat{r})^2}{2\sigma^2} + (1-\gamma)\bar{r} + \tfrac{1}{2}(1-\gamma)^2\sigma^2\right) d\bar{r}\, d\sigma^2 .$$

As function of σ^2 with $\bar{r}$ fixed, the exponential term in the integrand increases exponentially for large σ^2. The first term, $\sigma^{-2(m+1)}$, decreases in σ^2, but its polynomial rate of decrease is dominated by the exponential rate of increase of the exponential term, so the integrand does not go to zero as $\sigma^2 \to \infty$, and the integral is thus $+\infty$. Since this is true for ever $\bar{r}$, the joint integral is also infinite.

(c) (For those who want to do some of what MacKay calls "macho integration".) Compute, using numerical integration if necessary, the price of the asset assuming $C = 1$, $S = 1$, $\gamma = 2$, and that beliefs about $\bar{r}$ and σ^2 are arrived at from a prior joint pdf proportional to

$$\sigma^{-1} e^{-\frac{\bar{r}^2}{2\sigma^2} - \sigma^2}\, d\sigma^2\, d\bar{r} .$$

[This should make it possible to integrate $\bar{r}$ out of the posterior and the integrand in the pricing formula analytically, using the fact that the posterior will have a Normal shape as a function of $\bar{r}$. Then you probably have to use numerical integration with respect to σ^2.]

Will this prior imply a finite asset price for all positive values of γ?

I should have said that what I labeled as the "prior" pdf was the posterior after looking at the data, since I didn't specify any data.

Here we need to first find the constant of integration for the pdf of $\bar{r}, \sigma^2$, then integrate the pdf times the pricing formula. Note that the pdf of $\bar{r}, \sigma^2$ is $\sqrt{2\pi}$ times a $N(0, \sigma^2)$ for $\bar{r}$ conditional on σ^2, times an exponential pdf (the same as a $\gamma(1,1)$) in σ^2. Its integral is therefore just $\sqrt{2\pi}$. When we multiply the normalized pdf by the price conditional on $\bar{r}$ and σ^2 to take the expectation we get the integral

$$\int_0^\infty \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\bar{r}^2}{2\sigma^2} - \bar{r} + \tfrac{1}{2}\sigma^2 - \sigma^2\right) d\bar{r}\, d\sigma^2$$

Though this is not quite as obviously integrable as the pdf itself, note that the exponent is quadratic in $\bar{r}$ for fixed σ^2. That means it behaves as proportional to a normal. We need to "complete the square" for the quadratic form in $\bar{r}$. We rewrite the integral as

$$\int_0^\infty \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\bar{r}+\sigma^2)^2}{2\sigma^2}\right) d\bar{r}\, d\sigma^2$$

Unfortunately, it is now apparent that I miscalculated in setting up the problem. As a function of $\bar{r}$, this is a $N(-\sigma^2, \sigma^2)$ pdf. Therefore it integrates to one in $\bar{r}$ for each σ^2. And therefore when we then integrate with respect to σ^2, we get infinity. To avoid the infinite integral with $\gamma = 2$, I would have needed to make the marginall pdf of σ^2 $\lambda e^{-\lambda\sigma^2}$ with $\lambda > 1$. Or, with the stated prior, I could have made $\gamma < 2$. The general condition for the integral to be finite is $(1-\gamma)^2 < \lambda$.