Opinionated Lessons in Statistics

by Bill Press

#22 Uncertainty of Derived Parameters

What is the uncertainty in quantities other than the fitted coefficients:

Method 1: Linearized propagation of errors

nerdy math note: ∇f is technically a row (not column) vector, because it's a one-form

$\mathbf{b}_0$ is the MLE parameters estimate

$\mathbf{b}_1 \equiv \mathbf{b} - \mathbf{b}_0$ is the RV as the parameters fluctuate

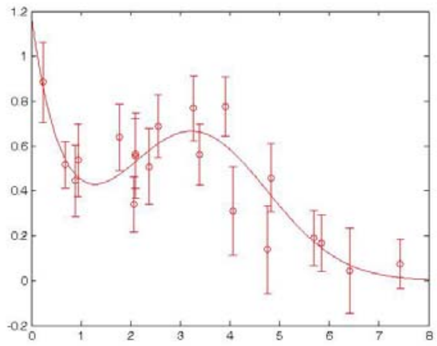
$$f \equiv f(\mathbf{b}) = f(\mathbf{b}_0) + \nabla f\ \mathbf{b}_1 + \cdots$$

$$\langle f \rangle \approx \langle f(\mathbf{b}_0) \rangle + \nabla f\ \cancel{\langle \mathbf{b}_1 \rangle} = f(\mathbf{b}_0)$$

$$\begin{aligned}\left\langle f^2 \right\rangle - \langle f \rangle^2 &\approx 2f(\mathbf{b}_0)(\nabla f\ \cancel{\langle \mathbf{b}_1 \rangle}) + \left\langle (\nabla f\ \mathbf{b}_1)^2 \right\rangle \\ &= \nabla f\ \left\langle \mathbf{b}_1 \mathbf{b}_1^T \right\rangle \nabla f^{\,T} \\ &= \nabla f\ \mathbf{\Sigma}_b \nabla f^{\,T}\end{aligned}$$

In our example, if we are interested in the area of the "hump",

```
bfit =
    1.1235    1.5210    0.6582    3.2654    1.4832
covar =
    0.1349    0.2224    0.0068   -0.0309    0.0135
    0.2224    0.6918    0.0052   -0.1598    0.1585
    0.0068    0.0052    0.0049    0.0016   -0.0094
   -0.0309   -0.1598    0.0016    0.0746   -0.0444
    0.0135    0.1585   -0.0094   -0.0444    0.0948
```



$$f = b_3 b_5$$

$$\nabla f = (0, 0, b_5, 0, b_3)$$

$$\nabla f \, \mathbf{\Sigma} \, \nabla f^T = b_5^2 \Sigma_{33} + 2 b_3 b_5 \Sigma_{35} + b_3^2 \Sigma_{55} = 0.0336$$

$$\sqrt{0.0336} = 0.18$$

So $b_3 b_5 = 0.98 \pm 0.18$ ← the one standard deviation (1-σ) error bar

Is it normally distributed?

Absolutely not! A function of normals is not normal (although, if they are all narrow, it might be close).

What is the uncertainty in quantities other than the fitted coefficients:

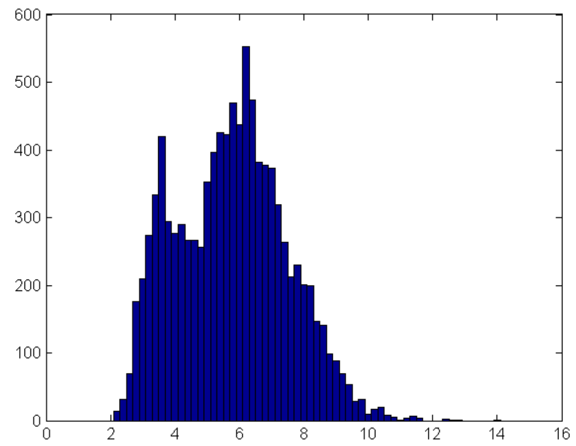
Method 2: Sample from the posterior distribution

1. Generate a large number of (vector) **b**'s

$$\mathbf{b} \sim \mathrm{MVNormal}(\mathbf{b}_0, \mathbf{\Sigma}_b)$$

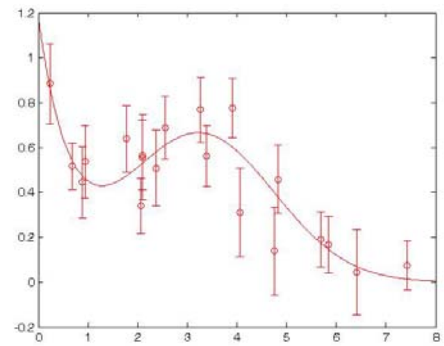
2. Compute your $f(\mathbf{b})$ separately for each **b**

3. Histogram



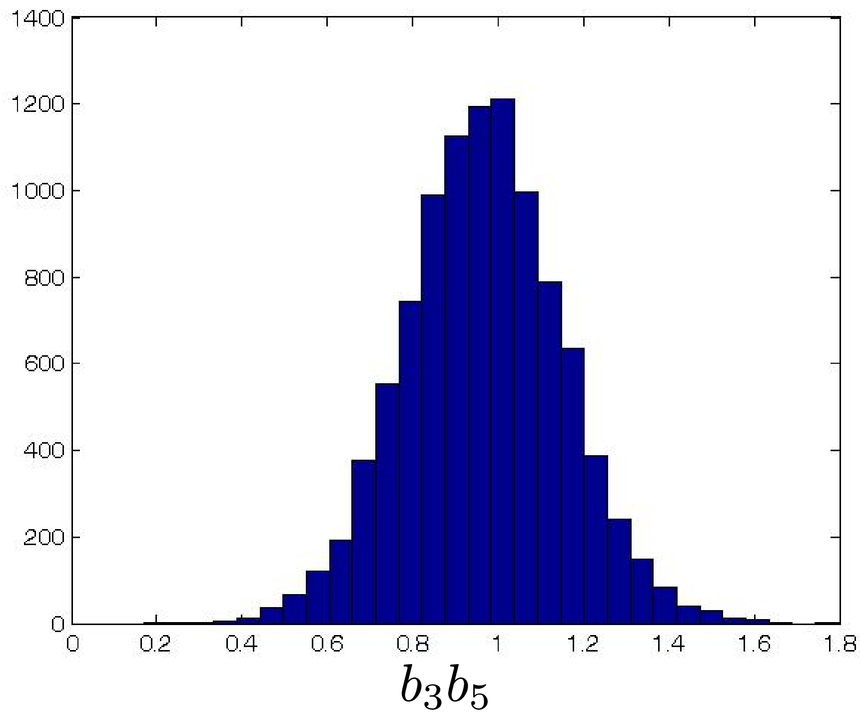
Note again that **b** is typically (close to) m.v. normal because of the CLT, but your (nonlinear) *f* may not, in general, be anything even close to normal!

Our example:




```
bees = mvnrnd(bfit,covar,10000);
humps = bees(:,3).*bees(:,5);
hist(humps,30);
std(humps)
```

std = 0.1833



Did I really need to use the full covar, not just the 2x2 piece for parameters 3 and 5?

Compare linear propagation of errors to sampling the posterior

- Note that even with lots of data, so that the distribution of the b's really → multivariate normal, a derived quantity might be very non-Normal.
  - In this case, sampling the posterior is a good idea!
- For example, the ratio of two normals of zero mean is Cauchy
  - which is very non-Normal!.
- So, sampling the posterior is a more powerful method than linear propagation of errors.
  - even when optimistically (or in ignorance) assuming multivariate Gaussian for the fitted parameters
- In fact, sampling the posterior distribution of large Bayesian models whose parameters are not at all Gaussian is, under the name MCMC, <u>the</u> most powerful technique in modern computational statistics.
  - we'll come back to this!