Bayes Rule and Updates in Beliefs

- From the previous slide:

$$P\left(\beta_{it}|\beta_{it}^{\text{realized}}\right) \propto P\left(\beta_{it}^{\text{realized}}|\beta_{it}\right) P\left(\beta_{it}\right).$$

Let's do some math!

- A preliminary calculation. Start with two known distributions:

$$\begin{aligned} x &\sim N\left(\mu_X, \sigma_X^2\right) \\ y|x &\sim N\left(x, \sigma_{Y|X}^2\right) \end{aligned}$$

- Here x corresponds to β_{it} and y corresponds to $\beta_{it}^{\text{realized}} = 1.8$
  - Further: μ_X is 1, $\sigma_X^2 = 0.5^2$, $\sigma_{Y|X}^2 = 0.4^2$ (the standard error squared of the realized beta)

Bayes Rule and Updates in Beliefs

- We want to get to the distribution of $x|y$, (really, $\beta_{it}|\beta_{it}^{\text{realized}}$), so let's first multiply these two pdf's:

$$\frac{1}{\sqrt{2\pi\sigma_X^2}}\exp\left\{\frac{\left(x-\mu_X\right)^2}{2\sigma_X^2}\right\}\frac{1}{\sqrt{2\pi\sigma_{Y|X}^2}}\exp\left\{\frac{\left(y-x\right)^2}{2\sigma_{Y|X}^2}\right\}$$

$$= \quad \frac{1}{2\pi\sqrt{\sigma_X^2\sigma_{Y|X}^2}}\exp\left\{\frac{\left(x-\mu_X\right)^2}{2\sigma_X^2}+\frac{\left(y-x\right)^2}{2\sigma_{Y|X}^2}\right\}$$

$$= \quad \frac{1}{2\pi\sqrt{\sigma_X^2\sigma_{Y|X}^2}}\exp\left\{\begin{array}{c}\frac{(x-\mu_X)^2}{2\sigma_X^2}\frac{\left(\sigma_X^{-2}+\sigma_{Y|X}^{-2}\right)^{-1}/\sigma_X^2}{\left(\sigma_X^{-2}+\sigma_{Y|X}^{-2}\right)^{-1}/\sigma_X^2}\ldots\\ +\frac{(y-x)^2}{2\sigma_{Y|X}^2}\frac{\left(\sigma_X^{-2}+\sigma_{Y|X}^{-2}\right)^{-1}/\sigma_{Y|X}^2}{\left(\sigma_X^{-2}+\sigma_{Y|X}^{-2}\right)^{-1}/\sigma_{Y|X}^2}\end{array}\right\}$$

Oh yeah... Algebra!

Continuing...

Define $k \equiv \left(\sigma_X^{-2} + \sigma_{Y|X}^{-2}\right)^{-1}$:

$$
\begin{aligned}
& \frac{1}{2\pi\sqrt{\sigma_X^2 \sigma_{Y|X}^2}} \exp\left\{ \frac{(x-\mu_X)^2}{2\sigma_X^2} \frac{k/\sigma_X^2}{k/\sigma_X^2} + \frac{(y-x)^2}{2\sigma_{Y|X}^2} \frac{k/\sigma_{Y|X}^2}{k/\sigma_{Y|X}^2} \right\} \\
= \quad & \frac{1}{2\pi\sqrt{\sigma_X^2 \sigma_{Y|X}^2}} \exp\left\{ \frac{\left(x^2 - 2x\mu_X + \mu_X^2\right) k/\sigma_X^2 + \left(y^2 - 2yx + x^2\right) k/\sigma_{Y|X}^2}{2k} \right\} \\
= \quad & \frac{1}{2\pi\sqrt{\sigma_X^2 \sigma_{Y|X}^2}} \exp\left\{ \frac{x^2 k/\sigma_X^2 - 2x\mu_X k/\sigma_X^2 + \mu_X^2 k/\sigma_X^2 + y^2 k/\sigma_{Y|X}^2 - 2yxk/\sigma_{Y|X}^2 + x^2 k/\sigma_{Y|X}^2}{2k} \right\} \\
= \quad & \frac{1}{2\pi\sqrt{\sigma_X^2 \sigma_{Y|X}^2}} \exp\left\{ \frac{x^2 k \left(\sigma_X^{-2} + \sigma_{Y|X}^{-2}\right) - 2x\left(yk/\sigma_{Y|X}^2 + \mu_X k/\sigma_X^2\right) + \mu_X^2 k/\sigma_X^2 + y^2 k/\sigma_{Y|X}^2}{2k} \right\}
\end{aligned}
$$

Finger-lickin'!

Continuing... ...

Note that $k\left(\sigma_X^{-2}+\sigma_{Y|X}^{-2}\right)=1$. So:

$$\frac{1}{2\pi\sqrt{\sigma_X^2\sigma_{Y|X}^2}}\exp\left\{\frac{x^2-2x\left(yk/\sigma_{Y|X}^2+\mu_X k/\sigma_X^2\right)+\mu_X^2 k/\sigma_X^2+y^2k/\sigma_{Y|X}^2}{2k}\right\}.$$

Next, complete the square:

$$\frac{1}{\sqrt{2\pi k}}\exp\left\{\frac{\left(x-\left(yk/\sigma_{Y|X}^2+\mu_X k/\sigma_X^2\right)\right)^2}{2k}\right\}$$

$$\times\frac{\sqrt{2\pi k}}{2\pi\sqrt{\sigma_X^2\sigma_{Y|X}^2}}\exp\left\{\frac{-\left(yk/\sigma_{Y|X}^2+\mu_X k/\sigma_X^2\right)^2+\mu_X^2k/\sigma_X^2+y^2k/\sigma_{Y|X}^2}{2k}\right\}.$$

Note that the first line says $x|y$ is normally distributed with mean $\left(yk/\sigma_{Y|X}^2+\mu_X k/\sigma_X^2\right)$ and variance k.

- The second line is a constant (not a function of x), conditional on y. Since we only were given the distribution up to a proportion (recall the Bayes Rule equation), we can ignore it for our purposes.

Learning with Normal Distributions

In sum, we are looking for the distribution of x conditional on a data point, y.

We found that $x|y$ is normally distributed using Bayes Rule.

- The mean of this distribution is:

$$
\begin{aligned}
yk/\sigma^2_{Y|X} + \mu_X k/\sigma^2_X &= y\frac{\sigma^{-2}_{Y|X}}{\sigma^{-2}_X + \sigma^{-2}_{Y|X}} + \mu_X \frac{\sigma^{-2}_X}{\sigma^{-2}_X + \sigma^{-2}_{Y|X}} \\
&= y \times (1 \text{ - weight on prior}) + \mu_X \times (\text{weight on prior})
\end{aligned}
$$

  Note that the more precise the signal is (the higher $\sigma^{-2}_{Y|X}$ is) and the less precise the prior is (the lower σ^{-2}_X is), the more weight is given to the signal when updating the mean belief about x.

- The variance is $k = \left(\sigma^{-2}_X + \sigma^{-2}_{Y|X}\right)^{-1} < \sigma^2_X$.