Value of Information

Henrik S. Lie

October 2019

Part I

a)

We assume that the profit of a project has a univariate Gaussian pdf,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right), \quad -\infty < x < \infty,$$

By a transformation of variables, $z = g(x) = \frac{x-\mu}{\sigma}, x = g^{-1}(z) = \mu + \sigma z$, we get that

$$p(z) = \left| \frac{dg^{-1}}{dz} \right| p(g^{-1}(z)) = \sigma \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\mu + \sigma z - \mu)^2}{2\sigma^2} \right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right),$$

which shows that z is standard Gaussian distributed.

b)

Then, we show that

$$\int_a^b z\phi(z)dz = \phi(a) - \phi(b),$$

by substituting $u = z^2/2, du = zdz$ into the integral, obtaining

$$\int_a^b z\phi(z)dz = \int_a^b \exp(-u)du = -\exp(-u)\Big|_a^b = -\phi(z)\Big|_a^b = \phi(a) - \phi(b).$$

c)

Now we suppose a decision maker is considering getting perfect information about the profit. The posterior value of perfect information is

$$PoV(x) = \int \max(0, x)p(x)dx = \int_0^\infty xp(x)dx.$$

Then we use the substitution from a), $z = g(x) = \frac{x-\mu}{\sigma}, dz = dx/\sigma$, and get

$$\int_{-\mu/\sigma}^\infty (\mu + \sigma z) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right) dz = \int_{-\mu/\sigma}^\infty \mu\phi(z)dz + \int_{-\mu/\sigma}^\infty \sigma z\phi(z)dz.$$

The first part is because of symmetry,

$$\int_{-\mu/\sigma}^{\infty} \mu\phi(z)dz = 1 - \mu\Phi(-\mu/\sigma) = \mu\Phi(\mu/\sigma),$$

and the second part is according to b), symmetry of $\phi(z)$ and the fact that $\lim_{z\to\infty} \phi(z) = 0$,

$$\int_{-\mu/\sigma}^{\infty} \sigma z\phi(z) = \sigma\phi(-\mu/\sigma) - \sigma\phi(\infty) = \sigma\phi(\mu/\sigma).$$

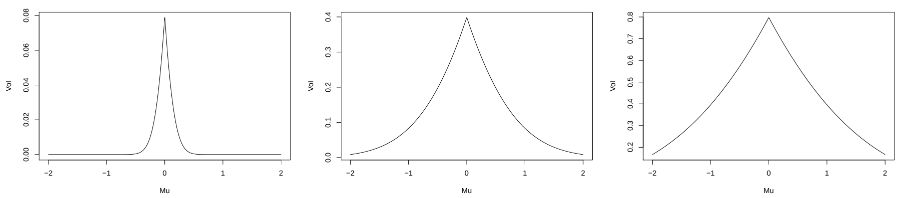
In total, we get,

$$PoV(x) = \mu\Phi(\mu/\sigma) + \sigma\phi(\mu/\sigma). \tag{1}$$

The prior value is $PV = \mu$, so we have $VoI = \mu\Phi(\mu/\sigma) + \sigma\phi(\mu/\sigma) - \mu$.

d)

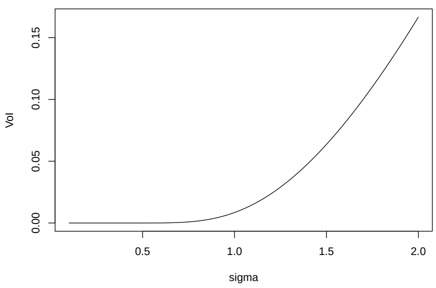
First, we plot the Value of Information (VoI) as a function of μ for different values of σ in Figure 1. From this figure, we see that the value of information is higher when the variability (σ) of the project outcome is higher. This is as expected, because with higher variability in outcome, it is much more valuable to know the actual outcome. Further, we see that the VoI is largest for values of μ around zero, and decreasing on $|\mu|$. The interpretation of this is as follows: for large absolute values of μ, the decision is simple and we know with high probability if the project will be profitable or not. Therefore, there is little value gained in getting to know the project outcome. This is not the case when μ is close to zero. For $\mu = 0$ there is a 50% chance of both positive and negative project outcome, so to then know the true outcome is highly valuable.



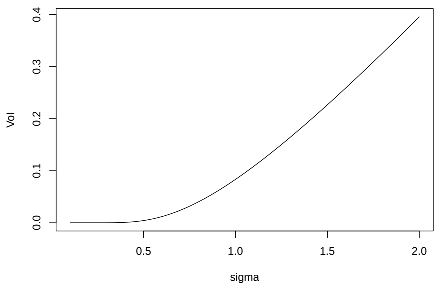
(a) $\sigma = 0.2$ (b) $\sigma = 1$ (c) $\sigma = 2$

Figure 1: VoI as function of μ for different values of σ.

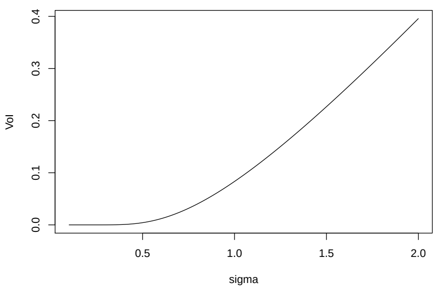
Next, we plot the VoI as a function of σ for different values of μ in Figure 2. First of all, we see that the dependency on σ is symmetric about μ. This means that the VoI is equal as a function of σ for positive and negative μ of equal absolute value.



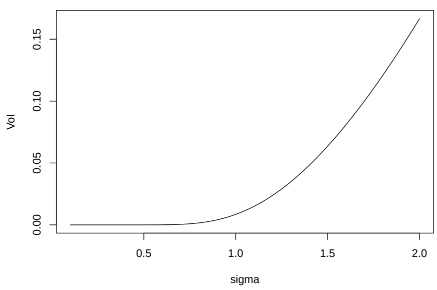
(a) $\mu = -2$



(b) $\mu = -1$



(c) $\mu = 1$



(d) $\mu = 2$

Figure 2: VoI as function of σ for different values of μ.

Finally, we plot a 2D plot of the VoI with μ on the x-axis and σ on the y-axis. From the plot in Figure 3, we see that VoI is largest for $\mu = 0$ and $\sigma = 2$.

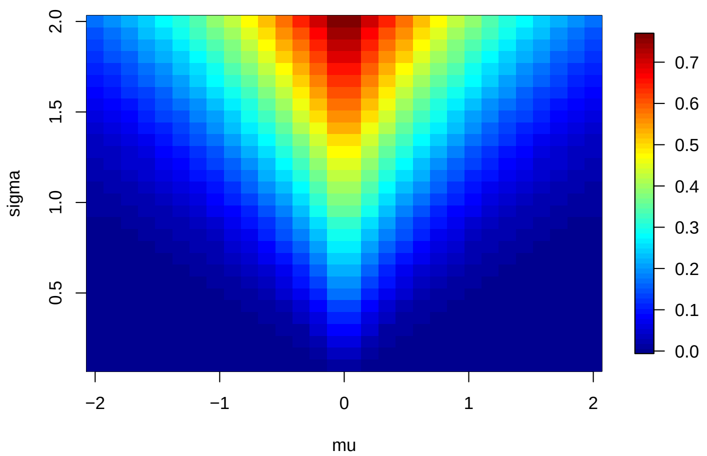


Figure 3

Part II

a)

For Gaussian random variables, we know that if $\mathbf{x}$ is Gaussian, then $\mathbf{y} = A\mathbf{x}+\mathbf{b}$ will also be Gaussian with mean $\boldsymbol{\mu}_y = A\boldsymbol{\mu}_x + \mathbf{b}$ and covariance matrix $\boldsymbol{\Sigma}_y = A\boldsymbol{\Sigma}_x A^T$. Then, we note that $\mathbf{y} = \mathbf{x}+\mathbf{b}$, with $\mathbf{b} \sim N(\mathbf{0}, \mathbf{T}), \mathbf{T} = \tau^2 I$. Therefore, we get $\mathbf{y} \sim p(\mathbf{y}) = N(\mathbf{0}, \boldsymbol{\Sigma} + \mathbf{T})$, because $\mathbf{x} \sim p(\mathbf{x}) = N(\mathbf{0}, \boldsymbol{\Sigma})$.

Following along the same lines, we get

$$E(\boldsymbol{\mu}_{x|y}) = \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}E(\mathbf{y}) = \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\mathbf{0} = \mathbf{0},$$

and

$$\begin{aligned}\mathrm{Var}(\boldsymbol{\mu}_{x|y}) &= \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\mathrm{Var}(\mathbf{y})\big(\boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\big)^T \\ &= \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}(\boldsymbol{\Sigma} + \mathbf{T})(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\boldsymbol{\Sigma}^T \\ &= \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\boldsymbol{\Sigma},\end{aligned}$$

because $\boldsymbol{\Sigma}$ and $\mathbf{T}$ both are symmetric. This means that we have $p(\boldsymbol{\mu}_{x|y}) = N(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\boldsymbol{\Sigma})$.

b)

Now, we want to use the previous results to compute the posterior value of total imperfect information

$$PoV(\mathbf{y}) = \sum_{j=1}^{2}\int \max(0, \mu_{\mathbf{x}|\mathbf{y},j})p(\mathbf{y})d\mathbf{y}.$$

Because only $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}}$ is informative for the decision, the expression simplifies to

$$PoV(\mathbf{y}) = \sum_{j=1}^{2}\int \max(0, \mu_{\mathbf{x}|\mathbf{y},j})p(\mu_{\mathbf{x}|\mathbf{y},j})d\mu_{\mathbf{x}|\mathbf{y},j}.$$

To compute this, we denote the variance of $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}}$ as

$$Var(\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}}) = \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{T})^{-1}\boldsymbol{\Sigma} = \mathbf{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{pmatrix}.$$

By properties of the multivariate Gaussian distribution, this means that $p(\mu_{\mathbf{x}|\mathbf{y},j}) = N(0, c_{jj})$. Therefore, we get

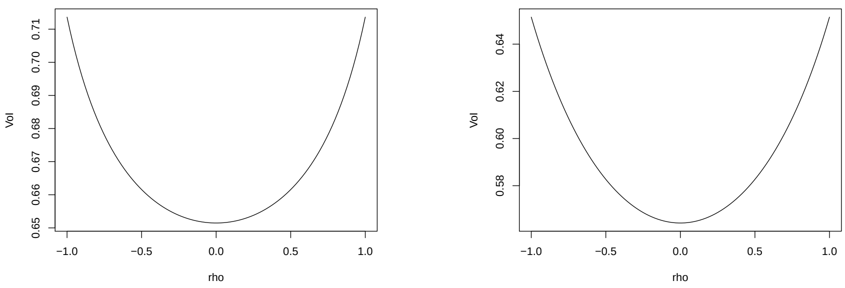
$$PoV(\mathbf{y}) = \sum_{j=1}^{2}\int_0^\infty \mu_{\mathbf{x}|\mathbf{y},j}p(\mu_{\mathbf{x}|\mathbf{y},j})d\mu_{\mathbf{x}|\mathbf{y},j}.$$

By the result in equation (1), we get

$$PoV(\mathbf{y}) = \sum_{j=1}^{2} 0\cdot\Phi(0/\sqrt{c_{jj}}) + \sqrt{c_{jj}}\phi(0/\sqrt{c_{jj}}) = \frac{\sqrt{c_{11}} + \sqrt{c_{22}}}{\sqrt{2\pi}}.$$

c)

In this exercise, we plot the value of information as a function of ρ for $\tau = 0.5$ and $\tau = 1$. The resulting graphs can be found in Figure 4. We see that the VoI is increasing for increasing values of $|\rho|$. Additionally, VoI is smaller for larger values of τ, which makes sense because larger τ implies less accurate information.



(a) $\tau = 0.5$ (b) $\tau = 1$

Figure 4: VoI as function of ρ for different values of τ.

d)

We now assume that the decision maker can gather information about only one of the projects, and we have $y_1 = \mathbf{F}\mathbf{x} + N(0, \tau^2)$, with $\mathbf{F} = (1, 0)$. Therefore, we get $y_1 \sim p(y_1) = N(0, \mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)$, because $\mathbf{x} \sim p(\mathbf{x}) = N(\mathbf{0}, \mathbf{\Sigma})$. Further, we have $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}} = \mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}y_1$

We follow along the same lines as in exercise a) and get

$$E(\boldsymbol{\mu}_{\mathbf{x}|y_1}) = \mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}E(y_1) = \mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}0 = \mathbf{0},$$

and

$$\begin{aligned}
\mathrm{Var}(\boldsymbol{\mu}_{\mathbf{x}|y_1}) &= \mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}\mathrm{Var}(y_1)\big(\mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}\big)^T \\
&= \mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}\mathbf{F}\mathbf{\Sigma}^T \\
&= \mathbf{\Sigma}\mathbf{F}^T(\mathbf{F}\mathbf{\Sigma}\mathbf{F}^T + \tau^2)^{-1}\mathbf{F}\mathbf{\Sigma},
\end{aligned}$$

because $\mathbf{\Sigma}$ is symmetric. Further, because

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

we get

$$\begin{aligned}
\mathrm{Var}(\boldsymbol{\mu}_{\mathbf{x}|y_1}) &= \mathbf{\Sigma}\mathbf{F}^T(1 + \tau^2)^{-1}\mathbf{F}\mathbf{\Sigma} \\
&= (1 + \tau^2)^{-1}\begin{pmatrix} 1 \\ \rho \end{pmatrix}\begin{pmatrix} 1 & \rho \end{pmatrix} \\
&= \frac{1}{1 + \tau^2}\begin{pmatrix} 1 & \rho \\ \rho & \rho^2 \end{pmatrix} \\
&= \mathbf{S}.
\end{aligned}$$

This means that we have $\boldsymbol{\mu}_{\mathbf{x}|y_1} \sim p(\boldsymbol{\mu}_{\mathbf{x}|y_1}) = N(\mathbf{0}, \mathbf{S})$.

We compute the Value of Information as we did in exercise b), and arrive at the expression

$$PoV(y_1) = \frac{\sqrt{s_{11}} + \sqrt{s_{22}}}{\sqrt{2\pi}} = \frac{1 + |\rho|}{\sqrt{2\pi(1+\tau^2)}}.$$

Finally, we plot the VoI, which is equal to the PoV because $PV = \boldsymbol{\mu} = 0$, for $\tau = 0.5$ and $\tau = 1$ and $\rho \in \{-1, 1\}$ in Figure 5.

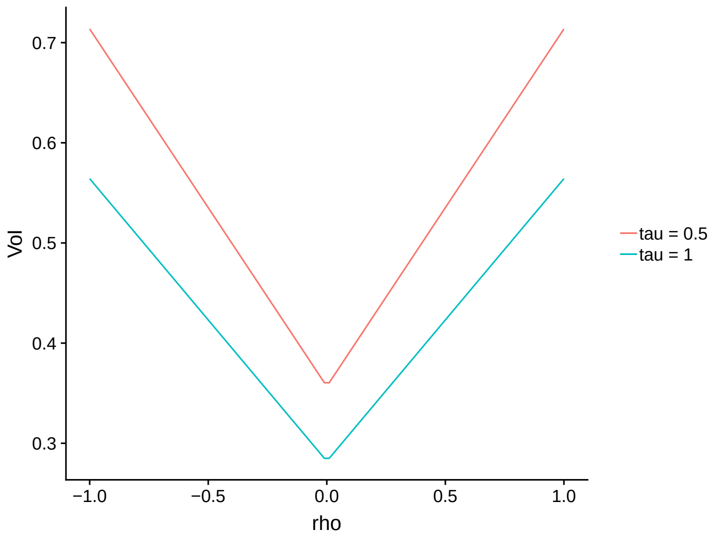


Figure 5: VoI as a function of ρ for two different values of τ.