VoI1

October 18, 2019

1 Value of Information Calculations for a Gaussian Example

1.1 Part I

We assume the profit of a project has a univariate Gaussain pdf,

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

The decision maker will invest in a project with a positive expected value.

1.1.1 a.

Given:

$$x = g^{-1}(z) = \mu + \sigma z$$

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

But: $z = g(x) = \frac{x-\mu}{\sigma}$

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

So this shows that the variable z is a standard notmal distributed with mean $\mu = 0$ and standard deviaiton $\sigma = 1$

■

1.1.2 b.

Show that:

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

Where: $\phi(z) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{z^2}{2}\right)$

$$\phi(a) - \phi(b) = \frac{1}{\sqrt{2\pi}}(\exp(\frac{-a^2}{2}) - \exp(\frac{-b^2}{2}))$$

Using integration by parts:

$$\int u \cdot v' = u \cdot v| - \int u' \cdot v$$

I could not do it with integration by parts, but with substitution as follows:

$$\int_a^b z\phi(z)dz = \int_a^b z\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)dz = I$$

let \$t = z^2 , dt = 2z dz \$ then: When $z \to a : t \to a^2$ When $z \to b : z \to b^2$

$$I = \int_{a^2}^{b^2} \frac{1}{\sqrt{2\pi}}\exp(\frac{-t}{2})\frac{dt}{2} = \frac{1}{2\sqrt{2\pi}}\int_{a^2}^{b^2}\exp(\frac{-t}{2})dt = \frac{1}{2\sqrt{2\pi}}\left|(-2\exp(\frac{-t}{2}))\right|_{a^2}^{b^2} = -\frac{2}{2\sqrt{2\pi}}(\exp(\frac{-b^2}{2}) - \exp(\frac{-a^2}{2})$$

■

Integration by parts:
If we set

$$v' = \frac{z \cdot \exp(\frac{-z^2}{2})}{\sqrt{2\pi}}$$
$$u = 1,$$

we get

$$v = -\frac{\exp(\frac{-z^2}{2})}{\sqrt{2\pi}} = -\phi(z)$$
$$u' = 0.$$

Thus, we obtain

$$\int_a^b z\phi(z)dz = \left| -\frac{\exp(\frac{-z^2}{2})}{\sqrt{2\pi}} \cdot 1\right|_{z=a}^{b} - \int_a^b 0 \cdot \phi(z)dz$$
$$= \left|\frac{\exp(\frac{-z^2}{2})}{\sqrt{2\pi}}\right|_{z=b}^{a} = \phi(a) - \phi(b)$$

1.1.3 c.

Assumne perfect information about the profit, then the posterior value of the perfect information value is:

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

$$\int_0^\infty xp(x)dx = \int_0^\infty \frac{(x-\mu+\mu)}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)dx$$

Substite for:

$$z = g(x) = \frac{x-\mu}{\sigma}$$

and

$$dx = \sigma dz$$

When $x \to 0 : z \to -\frac{\mu}{\sigma}$ When $x \to \infty : z \to \infty$

$$PoV = \int_0^\infty xp(x)dx = \int_{-\frac{\mu}{\sigma}}^\infty (z+\frac{\mu}{\sigma})\frac{1}{\sqrt{2\pi}}\exp(\frac{-z^2}{2})\sigma dz = \int_{-\frac{\mu}{\sigma}}^\infty (z\sigma+\mu)\phi(z)dz = \int_{-\frac{\mu}{\sigma}}^\infty z\sigma\phi(z)dz + \int_{-\frac{\mu}{\sigma}}^\infty \mu\phi(z)dz = I_1$$

The first part can be solve using the proof in section 1.b

$$I_1 = \sigma(\phi(-\frac{\mu}{\sigma}) - \phi(\infty)) = \sigma(\phi(-\frac{\mu}{\sigma})) = \sigma(\phi(\frac{\mu}{\sigma}))$$

The second integral is the standard form for Gaussian cdf Φ, and can be written as:

$$I_2 = \int_{-\frac{\mu}{\sigma}}^\infty \mu\phi(z)dz = \mu\int_{-\frac{\mu}{\sigma}}^\infty \phi(z)dz = \mu\Phi(z)|_{-\frac{\mu}{\sigma}}^\infty = \mu\Phi(\frac{\mu}{\sigma})$$

Subtite for I_1, I_2:

$$PoV(x) = \sigma\phi(\frac{\mu}{\sigma}) + \mu\Phi(\frac{\mu}{\sigma})$$

■

1.1.4 d.

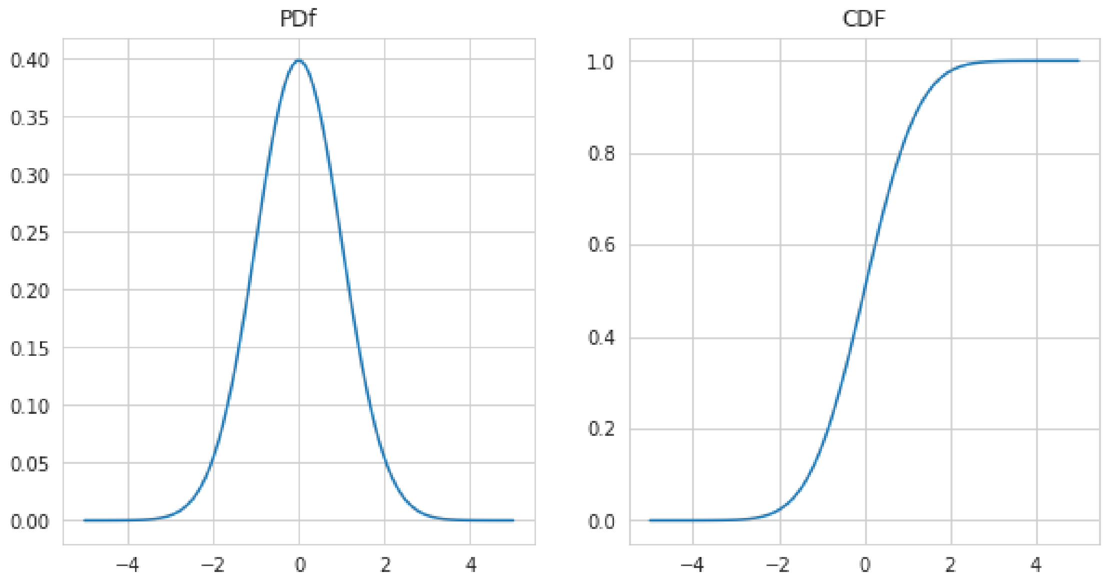
```
In [0]: import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        from scipy import integrate
        from scipy.stats import norm
```

```
In [0]: def VoI( mu , sigma):
            c = mu / sigma
            v1 = sigma * norm.pdf(c, 0 , 1)
            v2 = mu * norm.cdf(c, 0 , 1)
            return (v1 + v2 )
```

```
In [0]: z =  np.linspace(-5, 5, 100)
        z_pdf = np.zeros(100)
        z_cdf = np.zeros(100)
        for i in range(100):
            z_pdf[i] = norm.pdf(z[i], 0 , 1)
            z_cdf[i] = norm.cdf(z[i], 0 , 1)
```

```
In [0]: sns.set_style("whitegrid")

        fig, ax = plt.subplots(1,2)
        fig.set_size_inches(10, 5)
        ax[0].plot(z, z_pdf)
        ax[1].plot(z , z_cdf)
        ax[0].set_title("PDf")
        ax[1].set_title("CDF")
        plt.show()
```

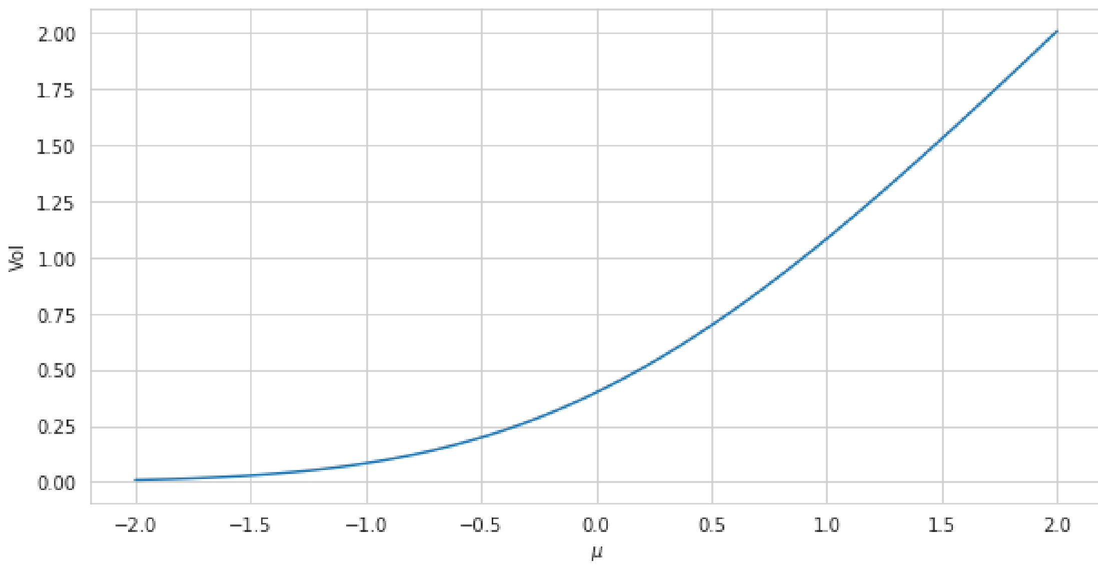


1. Plot of the analytical VoI as a function of the mean

```
In [0]: mu1 =  np.linspace(-2, 2, 100)
        sigma1 = 1
```

```
In [0]: VoI_mu = VoI(mu1 , sigma1)
```

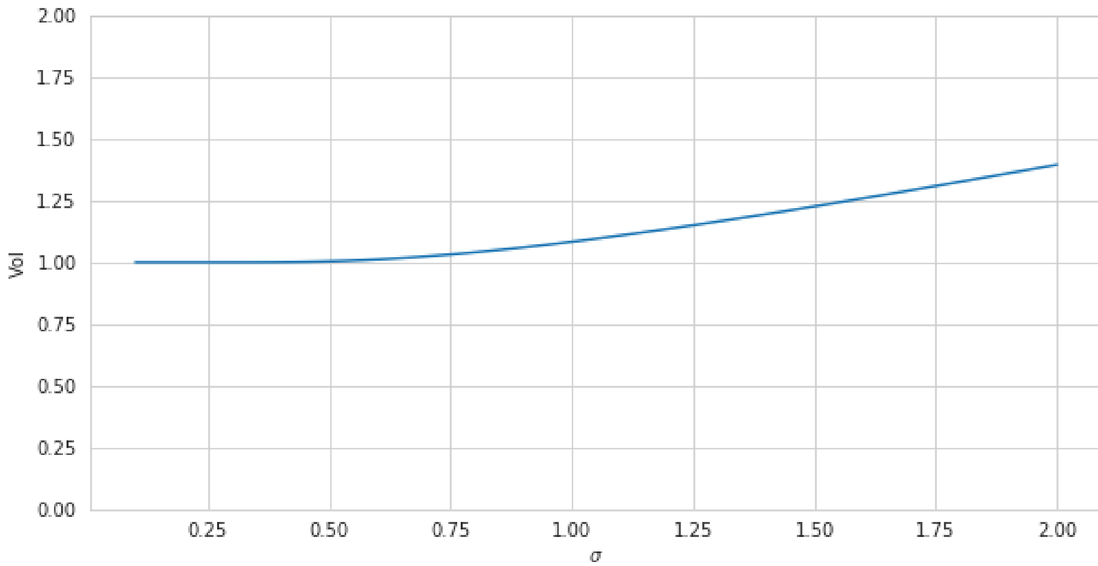
```
In [0]: sns.set_style("whitegrid")
        fig, ax = plt.subplots()
        fig.set_size_inches(10, 5)
        plt.plot(mu1 , VoI_mu)
        plt.ylabel("VoI")
        plt.xlabel("$\mu$")
        plt.show()
```



2. Plot of the analytical VoI as a function of the standard deviation

In [0]:
```
mu2 = 1
sigma2 =  np.linspace(0.1, 2, 100)
VoI_sigma = VoI(mu2 , sigma2)
```

In [0]:
```
sns.set_style("whitegrid")
fig, ax = plt.subplots()
fig.set_size_inches(10, 5)
plt.plot(sigma2 , VoI_sigma)
plt.ylabel("VoI")
plt.xlabel("$\sigma$")
plt.ylim(0 , 2)
plt.show()
```



From the figures above, we can notice that the VoI is more sensitives to changes in the mean value than the changes in the standard deviation.

■

1.2 Part II

In the second part of this exercise we will consider a multivariate Gaussian pdf

$$p(\boldsymbol{x}) = N(\mathbf{0}, \Sigma)$$

For profits

$$\boldsymbol{x} = (x_1, x_2)$$

and assuming the variance equal 1 for both projects and the correlation

$$-1 < \rho < 1.$$

1.2.1 a.

$$P(y) = P(\mathbf{x} + N(\mathbf{0}, \tau^2)) = P(x) + N(\mathbf{0}, \mathbf{T}) = N(\mathbf{0}, \mathbf{\Sigma}) + N(\mathbf{0}, \mathbf{T}) = N(\mathbf{0}, \mathbf{\Sigma} + \mathbf{T})$$

Where

$$T = \tau^2 I$$

■

$$P(x|y) \propto P(x)P(y|x) = N(0, \Sigma)N(x, T) = N(\mu_{x|y}, \Sigma_{x|y})$$

When two normal distributions are multiplied the resulting mean can be obtained as following:

$$\mu = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

$$\mu_{x|y} = \frac{\mathbf{0}\cdot\mathbf{T} + \mathbf{\Sigma}\cdot\mathbf{x}}{\mathbf{\Sigma}+\mathbf{T}} = \frac{\mathbf{\Sigma}\cdot\mathbf{x}}{\mathbf{\Sigma}+\mathbf{T}} = \mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{T})^{-\mathbf{1}}\cdot\mathbf{x}$$

$$P(\mu_{x|y}) = P(\mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{T})^{-\mathbf{1}}\cdot\mathbf{x}) = P(\mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{T})^{-\mathbf{1}})P(x) = N(\mathbf{0},\mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{T})^{-\mathbf{1}})N(\mathbf{0},\mathbf{\Sigma}) = N(\mathbf{0},\mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{T})^{-\mathbf{1}}\mathbf{\Sigma})$$

■

1.2.2 b.

In the case of the total imperfect information:

$$PoV(y) = \sum_{j=1}^{2}\int \max(0,\mu_{x_j|y_j})P(\mu_{x_j|y_j})d\mu_{x_j|y_j}$$

Let:

$$t_j = \mu_{x_j|y_j}$$

And

$$d\mu_{x_j|y_j} = dt_j$$

$$PoV(y) = \sum_{j=1}^{2}\int \max(0,t_j)P(t_j)dt_j = \sum_{j=1}^{2}\int t_jP(t_j)dt_j = \sum_{j=1}^{2} I_t$$

From the first part section c, we can write:

$$I_t = \sigma_t\phi(\frac{\mu_t}{\sigma_t}) + \mu_t\Phi(\frac{\mu_t}{\sigma_t})$$

$\mu_t = 0$, $\sigma_t^2 = \Sigma(\Sigma+T)^{-1}\Sigma$ Then $\sigma_t = \sqrt{\Sigma(\Sigma+T)^{-1}\Sigma}$

$$I_t = \sqrt{\Sigma(\Sigma+T)^{-1}\Sigma}\cdot\phi(0) + 0\cdot\Phi(0) = \sqrt{\Sigma(\Sigma+T)^{-1}\Sigma}\cdot(\frac{1}{\sqrt{(2\pi)}}) = \frac{\sqrt{\Sigma(\Sigma+T)^{-1}\Sigma}}{\sqrt{(2\pi)}}$$

Let:

$$S = \Sigma(\Sigma+T)^{-1}\Sigma$$

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

$$PoV(y) = \sum_{j=1}^{2}\frac{\sqrt{\Sigma(\Sigma+T)^{-1}\Sigma}}{\sqrt{(2\pi)}} = \sum_{j=1}^{2}\sqrt{\frac{\Sigma(\Sigma+T)^{-1}\Sigma}{2\pi}} = \frac{\sqrt{S_{1,1}}+\sqrt{S_{2,2}}}{\sqrt{2\pi}}$$

■

1.2.3 c.

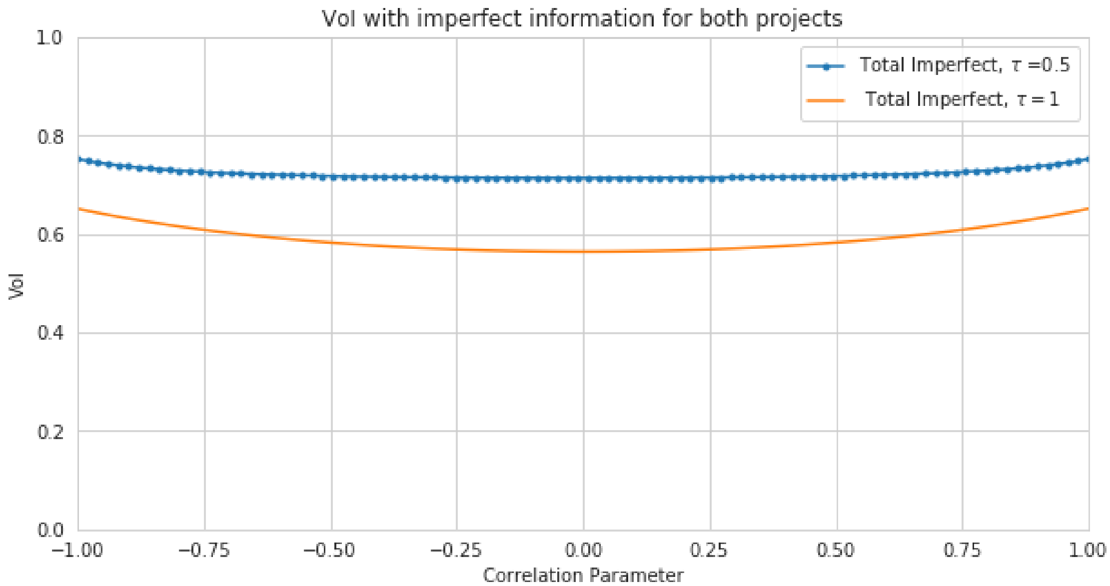
```
In [0]: tau_1 = 0.5
        tau_2 = 1
        rho = np.linspace(-1, 1, 100)
```

```
In [0]: def VoI_2(tau, r):
            c =np.sqrt(2 * np.pi)
            T =( tau ** 2)* np.identity(2)
            Sigma = np.array([[1,r],[r,1]])
            S = Sigma @ np.linalg.inv(Sigma + T) @ Sigma
            S11 = np.sqrt(S[0,0])
            S22 = np.sqrt(S[1,1])
            return (S11 + S22)/ c
```

```
In [0]: VoI_array1 = np.zeros(len(rho))
        VoI_array2 = np.zeros(len(rho))

        for i in range(len(rho)):
            VoI_array1[i]  = VoI_2(tau_1, rho[i])
            VoI_array2[i]  = VoI_2(tau_2, rho[i])
```

```
In [0]: sns.set_style("whitegrid")
        fig, ax = plt.subplots()
        fig.set_size_inches(10, 5)
        plt.plot(rho , VoI_array1 , marker = "." , label=r"Total Imperfect, $\tau$ =0.5")
        plt.plot(rho , VoI_array2 , label=r" Total Imperfect, $\tau = 1$")
        plt.ylabel("VoI")
        plt.xlabel("Correlation Parameter")
        plt.xlim(-1,1)
        plt.ylim(0, 1)
        plt.legend()
        plt.title("VoI with imperfect information for both projects")
        plt.show()
```



The above graph shows that the VOI slightly increase as the correlation between the two projects increase. At the same time, increasing the standard deviation result in decreasig the VoI. ■

1.2.4 d.

In the case of imperfect information about one project:

$$PoV(y) = \sum_{j=1}^{2} \int \max(0, \mu_{x_j|y_j}) P(\mu_{x_j|y_j}) d\mu_{x_j|y_j}$$

$$\mu_{x|y} = \Sigma \mathbf{F}^t (\mathbf{F} \Sigma \mathbf{F}^t + \tau^2)^{-1} y_1$$

Let:

$$t = \mu_{x_j|y_j}$$
$$d\mu_{x_j|y_j} = dt$$

$$PoV(y_1) = \sum_{j=1}^{2} \int \max(0, t_j) P(t_j) dt_j = \sum_{j=1}^{2} \int t_j P(t_j) dt_j = \sum_{j=1}^{2} I_t$$

From the first part section c, we can write:

$$I_t = \sigma_t \phi(\frac{\mu_t}{\sigma_t}) + \mu_t \Phi(\frac{\mu_t}{\sigma_t})$$

$\mu_t = 0$, $\sigma_t^2 = \Sigma \mathbf{F}^t (\mathbf{F} \Sigma \mathbf{F}^t + \tau^2)^{-1} \tau^2$ Then $\sigma_t = \sqrt{\Sigma \mathbf{F}^t (\mathbf{F} \Sigma \mathbf{F}^t + \tau^2)^{-1} \tau^2}$
Where: $\mathbf{F} = (1, 0)$ Let:

$$S = \Sigma \mathbf{F}^t (\mathbf{F} \Sigma \mathbf{F}^t + \tau^2)^{-1} \tau^2$$

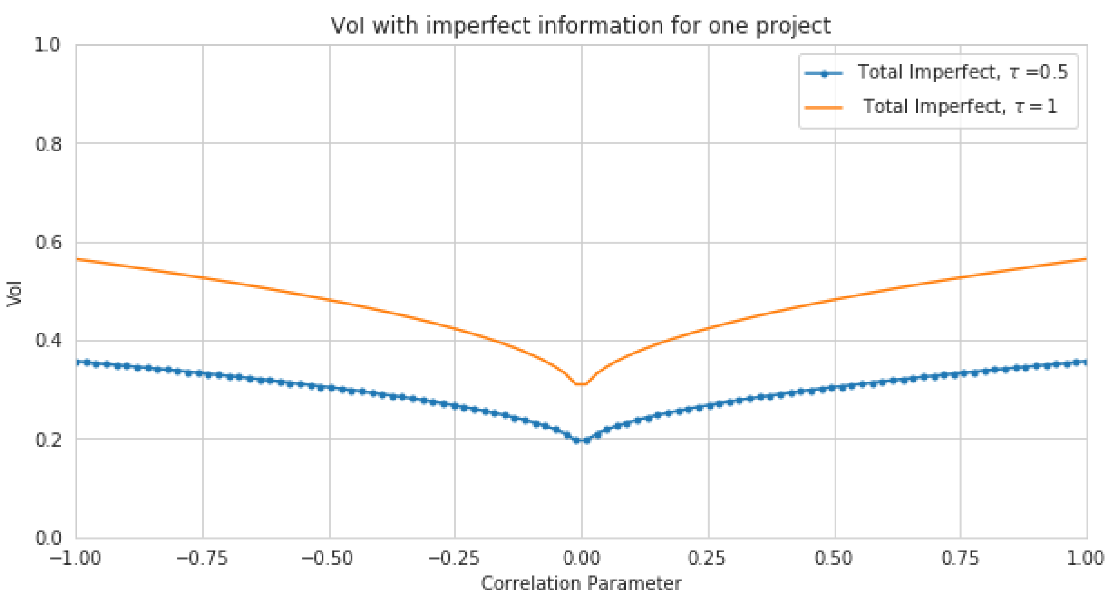
$$PoV(y) = \frac{\sqrt{|S_1|} + \sqrt{|S_2|}}{\sqrt{2\pi}}$$

```
In [0]: def VoI_3(tau, r, F):
            c =np.sqrt(2 * np.pi)
            Ft = F.transpose()
            Sigma = np.array([[1,r],[r,1]])
            T =( tau ** 2)* np.identity(1)
            S = (Sigma @ Ft) * 1 / (F @ Sigma @ Ft + T) * T
            S1 =  np.sqrt(abs (S[0]))
            S2 = abs (np.sqrt(abs( S[1])))
            return (S1 + S2)/ c
```

```
In [0]: F = np.array([1,0], ndmin = 2)
        VoI_array3 = np.zeros(len(rho))
        VoI_array4 = np.zeros(len(rho))

        for i in range(len(rho)):
            VoI_array3[i]  = VoI_3(tau_1, rho[i] , F)
            VoI_array4[i]  = VoI_3(tau_2, rho[i] , F)
```

```
In [0]: sns.set_style("whitegrid")
        fig, ax = plt.subplots()
        fig.set_size_inches(10, 5)
        plt.plot(rho , VoI_array3 , marker = "." , label=r"Total Imperfect, $\tau$ =0.5")
        plt.plot(rho , VoI_array4 , label=r" Total Imperfect, $\tau = 1$")
        plt.ylabel("VoI")
        plt.xlabel("Correlation Parameter")
        plt.xlim(-1,1)
        plt.ylim(0, 1)
        plt.legend()
        plt.title("VoI with imperfect information for one project")
        plt.show()
```



In the figure above we can see the effect of the correlation on the value of information when we can gather information from only one project from the two. VoI increase as the correlation increase.

■