# Solution to analysis in Home Assignment 1

Nicholas Granlund, nicgra@student.chalmers.se

## Analysis

In this report I will present my independent analysis of the questions related to home assignment 1. I have discussed the solution with NONE, NONE and NONE and I swear that the analysis written here are my own.

# 1 Properties of random variables

**a)** We start of with showing i):

We start of from this expression:

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} xp(x) \ dx$$

Expanding $p(x)$ we get:

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ dx$$

Factoring terms not dependent on x out from the integral gives

$$\mathbb{E}[x] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}} \ dx$$

Now we use the substitution $y = \frac{(x-\mu)}{\sigma}$ which gives us $x = \sigma y + \mu$ and $dx = \sigma dy$.

$$\mathbb{E}[x] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} (\sigma y + \mu)e^{-\frac{y^2}{2}} \ dy$$

Notice how the $\sigma^2$ in the denominator in the constant term vanishes. This is due to the substitution $dx = \sigma dy$, where the $\sigma$ cancels. Furthermore expanding and simplifying we get two integrals that are easier evaluated separately:

$$\mathbb{E}[x] = \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty} \sigma ye^{-\frac{y^2}{2}} \ dy + \int_{-\infty}^{\infty} \mu e^{-\frac{y^2}{2}} \ dy\right)$$

The first integral will evaluate to 0 since it is symmetric. The second integral will evaluate to $\mu\sqrt{2\pi}$ (This calculation was done with the help of Wolfram Alpha integral evaluator). This finally gives us

$$\mathbb{E}[x] = \frac{1}{\sqrt{2\pi}}\left(0 + \mu\sqrt{2\pi}\right)$$

$$\mathbb{E}[x] = \mu$$

Now we show ii);

We start of from the expression which defines variance:

$$\mathrm{Var}[x] = \mathbb{E}[(x-\mu)^2]$$

This can be expanded to:

$$\mathrm{Var}[x] = \mathbb{E}[x^2 - 2x\mu + \mu^2]$$

Using the results from i) that $\mathbb{E}[x] = \mu$ we get

$$\mathrm{Var}[x] = \mathbb{E}[x^2] - 2\mu\mathbb{E}[x] + \mu^2$$

Since $\mu$ is not a variable, but a constant/scalar. The expected value of this is simply the value itself. Therefore $\mathbb{E}[\mu] = \mu$ and we treat it as such in the expression above. Continuing, by using the definition of expected value again we can rewrite the expressions using the familiar integrals:

$$\text{Var}[x] = \int_{-\infty}^{\infty} x^2 p(x)\,dx - 2\mu^2 + \mu^2$$

Simplifying we get

$$\text{Var}[x] = \int_{-\infty}^{\infty} x^2 p(x)\,dx - \mu^2$$

Letting $y = \frac{(x-\mu)}{\sigma}$, we can rewrite the integral as:

$$\text{Var}[x] = \sigma^2 \int_{-\infty}^{\infty} (\sigma y + \mu)^2 p(y)\,dy - \mu^2$$

Expanding this we get

$$\text{Var}[x] = \sigma^2 \left( \int_{-\infty}^{\infty} y^2 p(y)\,dy + 2\mu \int_{-\infty}^{\infty} y p(y)\,dy + \mu^2 \int_{-\infty}^{\infty} p(y)\,dy - \mu^2 \right)$$

The first integral is 1 since it is normalized by $p(y)$, the second integral is just the expected value of $y$ which is 0. The third integral is the area under the probability density function which is 1. This gives

$$\text{Var}[x] = \sigma^2(1 + 0 + \mu^2 - \mu^2)$$
$$\text{Var}[x] = \sigma^2$$

**b)** We start of with showing i):

$$\mathbb{E}[z] = \int z\, p(z) \;\, dz$$
$$\mathbb{E}[z] = \int Aq\, p(q) \;\, dq$$
$$\mathbb{E}[z] = A \int q\, p(q) \;\, dq$$
$$\mathbb{E}[z] = A\mathbb{E}[q]$$

Now we show ii):

$$\mathrm{Cov}[z] = \mathrm{Cov}[Aq] = \mathbb{E}[(Aq - \mathbb{E}[Aq])(Aq - \mathbb{E}[Aq])^\top]$$
$$\mathrm{Cov}[z] = \mathbb{E}[(Aq - A\mathbb{E}[q])(Aq - A\mathbb{E}[q])^\top]$$
$$\mathrm{Cov}[z] = \mathbb{E}[A(q - \mathbb{E}[q])(q - \mathbb{E}[q])^\top A^\top]$$
$$\mathrm{Cov}[z] = A\, \underbrace{\mathbb{E}[(q - \mathbb{E}[q])(q - \mathbb{E}[q])^\top}_{\mathrm{Cov}(q)} A^\top]$$
$$\mathrm{Cov}[z] = A\mathrm{Cov}[q]A^\top$$

**c)**

The transformation results in the following:

$$p(q) = \mathcal{N}\Big(q; \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 0.3 & 0.0 \\ 0.0 & 8.0 \end{bmatrix}\Big)$$

$$p(z) = \mathcal{N}\Big(z; \begin{bmatrix} 5 \\ 10 \end{bmatrix}, \begin{bmatrix} 2.3 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}\Big)$$

The illustration of the mean and covariance can be seen in Figure 1.1 below. We can see that the transformation affects the mean and the covariance (as shown earlier in the coding part of the assignment). Since the transformation matrix $A$ is upper diagonal and has the diagonal elements of 1, the second components mean and variance remains unchanged. After the tranformation the mean of the second component is 10, same as before the transformation, furhtermore the variance of the second component is also unaltered (still 8). However, the transformation has introduced cross-covariance between the components (the off-diagonal elements for the covariance matrix are non zero). This implies that there is some correlation between the variance between the first and second component, which is illustrated by the $3\sigma$-curve being skewed right.

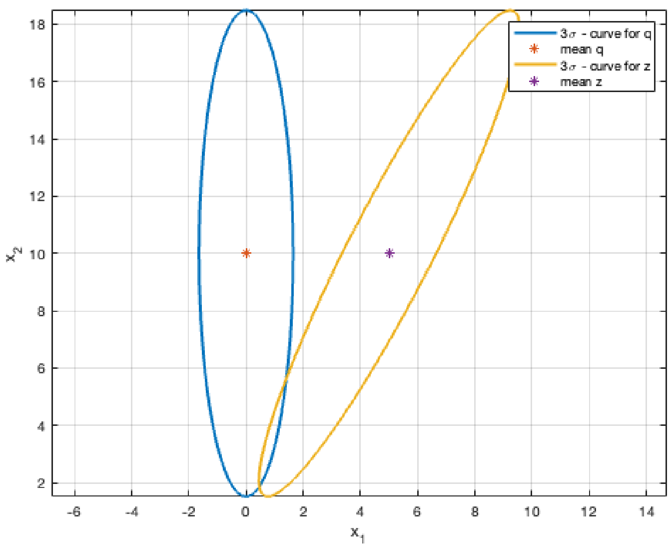


Figure 1.1: Illustration of the distribution of $q$ and $z$.

## 2 Transformation of random variables

**a)** By using the function **approxGaussianTransform** and **affineGaussianTransform** we are able to illustrate samples from the distributions along with the correspoinding PDF curve given the approximated mean and variance. The result can be seen in Figure 2.1 below. We use 5000 samples to approximate the PDF in both cases, and evidently they are rather similar. This is expected since it follows the law of large numbers, increasing the sample size make the sample mean and variance converge to the population mean and variance. By setting the sample size to 100 000 the approximated mean and variance are more or less spot on, and by setting the sample size to 10, the approximation is far worse. The result are as follows:

$$x \sim \mathcal{N}\Big(0, 2\Big)$$

$$\hat{z}_{analytical} \sim \mathcal{N}\Big(0, 18\Big)$$

$$\hat{z}_{approximate} \sim \mathcal{N}\Big(0.0276, 17.82808\Big)$$

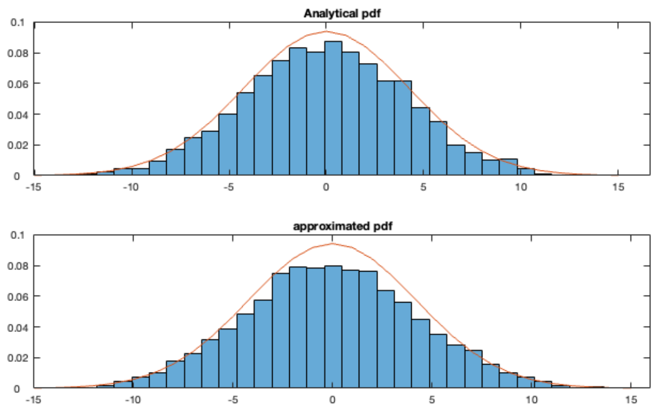


Figure 2.1: Analytical and approximated PDF. 5000 samples was used to approximate. They match rather well

**b)**

Since $z = x^3$ is not a linear transformation, we cannot obtain an analytical expression for the distribution mean and variance. However, we can still try to approximate it! By once again using the function **approxGaussianTranform** we can estimate the mean and variance of the distribution of $z$. the result can be seen in Figure 2.2 below. The result are as follows:

$$x \sim \mathcal{N}\Big(0, 2\Big)$$

$$\hat{z}_{analytical} \sim N/A$$

$$\hat{z}_{approximate} \sim \mathcal{N}\Big(0.0067, 120.7747\Big)$$

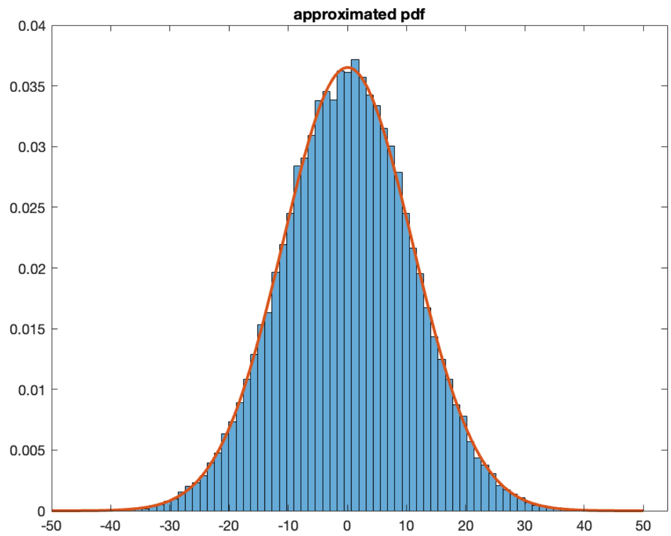


Figure 2.2: Approximated PDF. 50 000 samples was used to approximate. The histogram has 75 bins. Large sample size is needed due to the large variance subject to the non-linearity in the transformation.

**c)**

We can conclude that the non-linear transformation has a large influence on the variance of the transformed variable $z$. This specific non-linear transformation has no effect on the mean which is not always the case. E.g, if the transformation was $z = x^3 + b$, we would expect the mean to be shifted towards $b$. We can also conclude that there is no analytical solution to determine the mean and variance of a variable transformed through non-linear means.

# 3 Understanding the conditional density

**a)**

Yes, it is possible! The function $h(x)$ is deterministic and thus has no impact on the variance, but only the mean of $y$. We get $\mathbb{E}[y] = \mathbb{E}[h(x)] + \mathbb{E}[r] = h(x)$. One way to think about the deterministic function/variable $h(x)$, is as a Gaussian variable with mean $h(x)$ and variance 0 (no uncertainties whatsoever). It is therefor only a static mean. The variance of $y$ is hence only governed by $r$. To conclude we get $y \sim \mathcal{N}(h(x), \sigma_r^2)$.

**b)**

Yes, it is still possible. We do not obtain any new information given $x$ due to the deterministic nature of $h(x)$. Therefor the distribution of $y$ is $y \sim \mathcal{N}(h(x), \sigma_r^2)$, just like the previous task. the answer would be different if $h(x)$ was not deterministic but instead randomly distributed.

**c)**
**d)**
**e)**

# 4 MMSE and MAP estimators

**a)**

By sampling $\theta$ and $w$ we get the results in Figure 4.1. We notice that the distribution looks rather similar to a distribution described by $0.5p(\theta) + 0.5p(w)$, and sure enough, if we plot the distribution $p(y) = 0.5p(\theta) + 0.5p(w)$ we get a perfect fit.

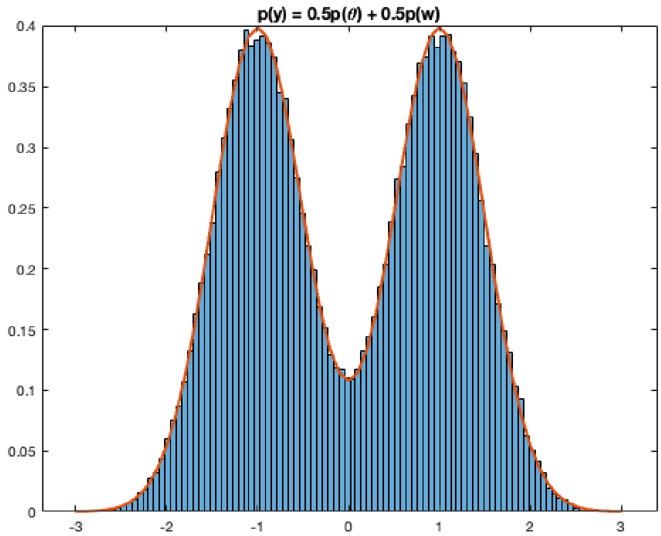


Figure 4.1: $y = \theta + w$ sampled 10 000 times. histogram of y plotted alongside $p(y)$

**b)**

If I were given the sample $y = 0.7$, my guess would be that $\theta = 1$. I would follow the logic of MAP, in which we find the $\theta$ that has the highest probability of yielding a measurement $y = 0.7$, this would be $\theta = 1$. With that beeing said, there is still a small probability that the actual $\theta = -1$ where the noise $w$ is exceptionally large.

**c)**

**d)**

**e)**

**f)**

**g)**