# Solution Discussion HA1

# Properties of random variables

a-i)

Since $x \sim \mathbb{N}(\mu, \sigma^2)$, the pdf is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

by definition of expectation

$$\begin{aligned}
\mathbb{E}(x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu) e^{-t^2} dt \\
&= \frac{1}{\sqrt{\pi}} \left( \sqrt{2}\sigma \int_{-\infty}^{\infty} t e^{-t^2} dt + \mu \int_{-\infty}^{\infty} e^{-t^2} dt \right) \\
&= \frac{1}{\sqrt{\pi}} \left( \sqrt{2}\sigma \left( -\frac{1}{2} e^{-t^2} \right) + \mu\sqrt{\pi} \right) \Bigg|_{t=-\infty}^{t=\infty} \\
&= \frac{\mu\sqrt{\pi}}{\sqrt{\pi}} \\
&= \mu
\end{aligned}$$

# Properties of random variables

a-ii)

$$
\begin{aligned}
\mathrm{Var}[x] &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx - \mu^2 \\
&= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu)^2 e^{-t^2}\,dt - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}\left(2\sigma^2\int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt + 2\sqrt{2}\sigma\mu\int_{-\infty}^{\infty} t e^{-t^2}\,dt + \mu^2\int_{-\infty}^{\infty} e^{-t^2}\,dt\right) - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}\left(2\sigma^2\int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt + 2\sqrt{2}\sigma\mu\left(-\frac{1}{2}e^{-t^2}\right) + \mu^2\sqrt{\pi}\right) - \mu^2\Bigg|_{t=-\infty}^{t=\infty} \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\left(-\frac{t}{2}e^{-t^2}\Bigg|_{t=-\infty}^{t=\infty} + \frac{1}{2}\int_{-\infty}^{\infty} e^{-t^2}\,dt\right) \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\int_{-\infty}^{\infty} e^{-t^2}\,dt \\
&= \frac{2\sigma^2\sqrt{\pi}}{2\sqrt{\pi}} \\
&= \sigma^2
\end{aligned}
$$

# Properties of random variables

b) z = Aq, and A is a constant matrix

i) $$\mathbf{z} = \mathbf{A}\mathbf{q} \implies \mathbb{E}(\mathbf{z}) = \mathbb{E}(\mathbf{A}\mathbf{q}) \implies \int_{-\infty}^{\infty} \mathbf{z} f(\mathbf{z}) dz = \mathbf{A} \int_{-\infty}^{\infty} \mathbf{q} f(\mathbf{q}) dq \implies \mathbb{E}(\mathbf{z}) = \mathbf{A}\mathbb{E}(\mathbf{q})$$

ii) $$Cov(\mathbf{z}) = \mathbb{E}[(\mathbf{z} - \mathbb{E}[\mathbf{z}])(\mathbf{z} - \mathbb{E}[\mathbf{z}])^T]$$

$$= Aq \;= \mathbb{E}[(\mathbf{A}\mathbf{q} - \mathbf{A}\mathbb{E}[\mathbf{q}])(\mathbf{A}\mathbf{q} - \mathbf{A}\mathbb{E}[\mathbf{q}])^T]$$

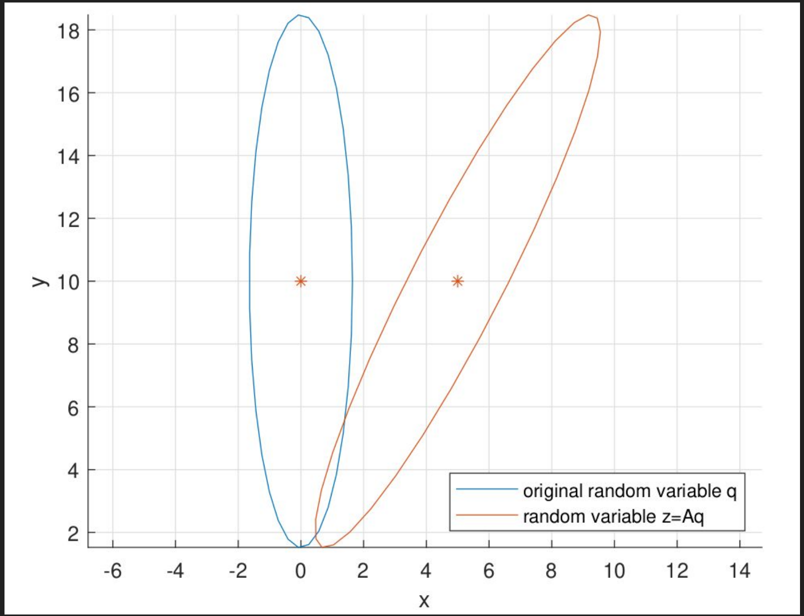
$$= \mathbb{E}[\mathbf{A}(\mathbf{q} - \mathbb{E}(\mathbf{q}))(\mathbf{q} - \mathbb{E}(\mathbf{q}))\mathbf{A}^T]$$

$$= \mathbf{A}\mathbb{E}[(\mathbf{q} - \mathbb{E}(\mathbf{q}))(\mathbf{q} - \mathbb{E}(\mathbf{q}))]\mathbf{A}^T = \mathbf{A}\, Cov(\mathbf{q})\, \mathbf{A}^T$$

# Properties of random variables

- $$p(\mathbf{q}) = \mathcal{N}\left(\mathbf{q}; \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 0.3 & 0 \\ 0 & 8 \end{bmatrix}\right) \quad \mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}.$$

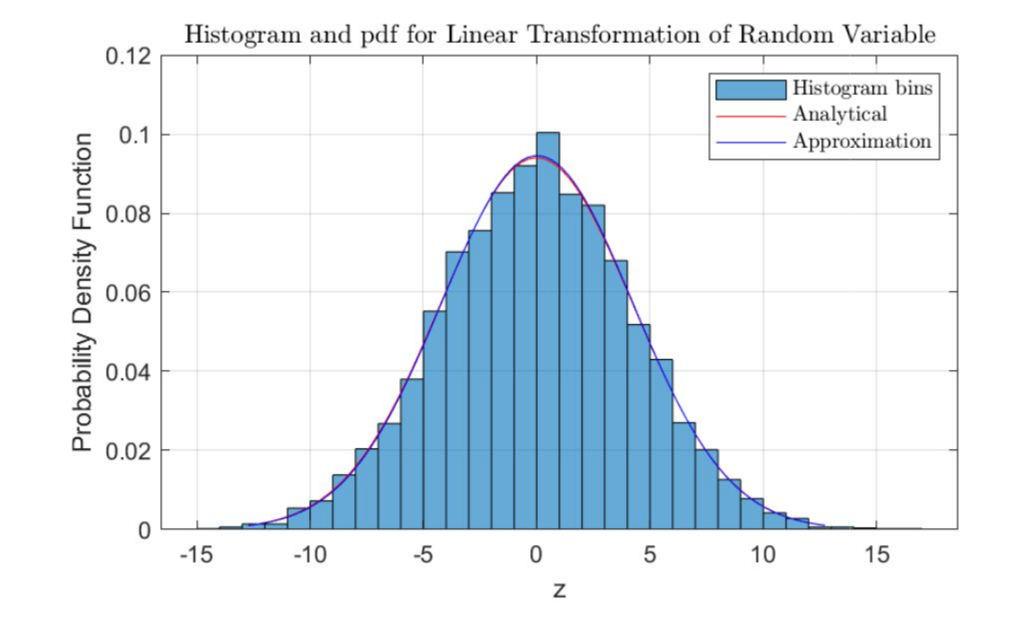
- The transformed x-coordinate is a linear combination of the original x-coordinate and the original y-coordinate with factors 1 and 0.5, respectively.



# Transformation of random variables

a) $\mathbb{E}[z] = \mathbb{E}[3x] = 3\mathbb{E}[x] = 3 \cdot 0 = 0$

$\mathrm{Cov}[z] = \mathrm{Cov}[3x] = 3 \cdot \mathrm{Cov}[x] \cdot 3 = 18$




- N = 5000
- More samples, fits better!

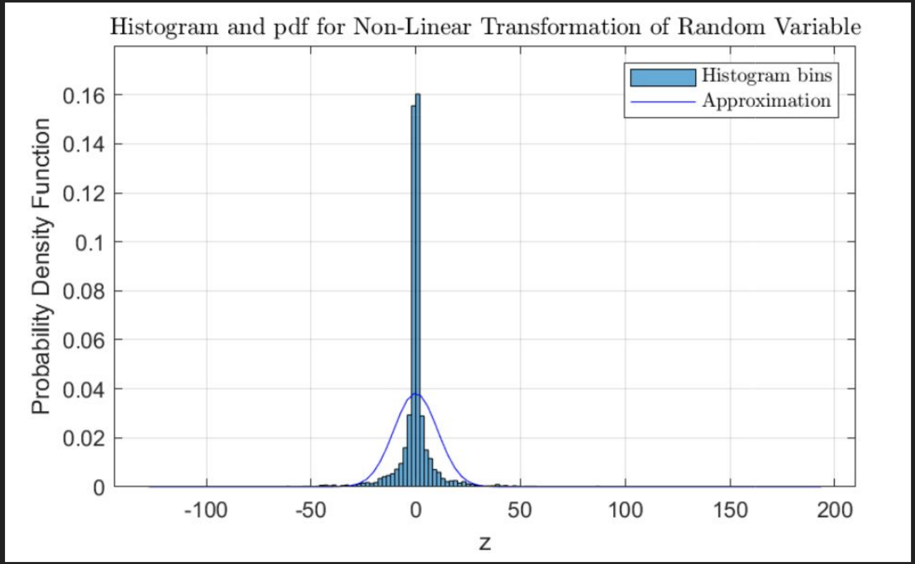
# Transformation of random variables

b) $z = x^3$

- Non-gaussian distribution
- Analytical mean and covariance can be calculated, but z does not follow a Gaussian distribution

c)

- linear transformation
- non-linear transformation



# Understanding the conditional density

a) Could be impossible→ p(x) is unknown, h(x) is unknown

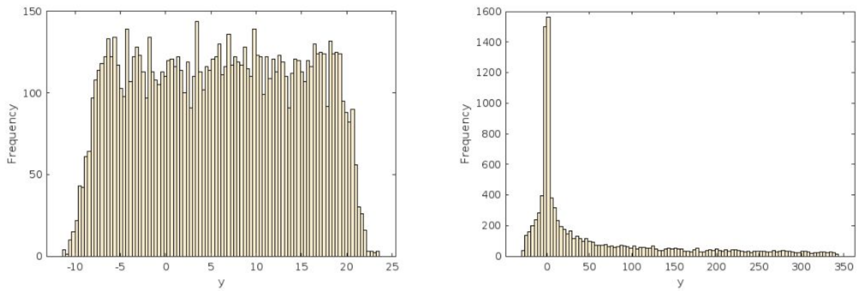
b) Yes, given x, h(x) is deterministic function, r is gaussian noise

$$p(y|x) = \mathcal{N}(h(x), \sigma^2)$$

c) h(x) =H(x), H is deterministic and constant
   i) p(y) → no
   ii) p(y|x) → yes

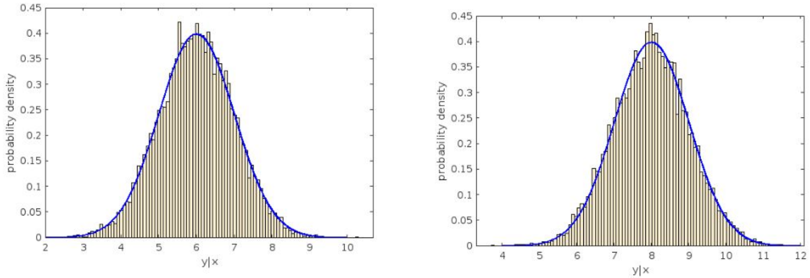
d) Given p(x) is Gaussian,
   i) p(y) → depends on the linearity of h(x)
      1) Linear → Gaussian
      2) non-linear → unknown
   ii) p(y|x) → yes, since h(x) is deterministic.

# Understanding the conditional density--Simulation



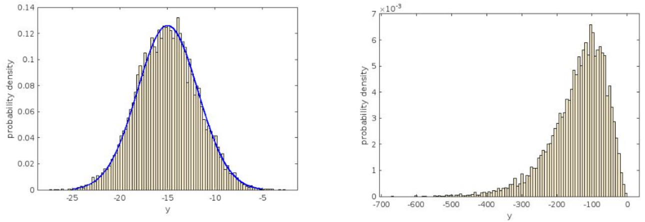

(a) $h(x) = 3x$. (b) $h(x) = x^3$.

Figure 3.1: Simulation for **a)**. In this simulation, $x \sim \text{Uniform}(-3, 7)$ and $r \sim \mathcal{N}(0, 1)$. Regardless of whether $h(x) = 3x$ or $x^3$, we cannot describe $p(y)$.




(a) $h(x) = 3x$. (b) $h(x) = x^3$.

Figure 3.2: Simulation for **b)**. In this simulation, $x = 2$ and $r \sim \mathcal{N}(0, 1)$. When $h(x) = 3x$, $y|x \sim \mathcal{N}(6, 1)$; when $h(x) = x^3$, $y|x \sim \mathcal{N}(8, 1)$. Regardless of whether $h(x) = 3x$ or $x^3$, we can describe $p(y|x)$.



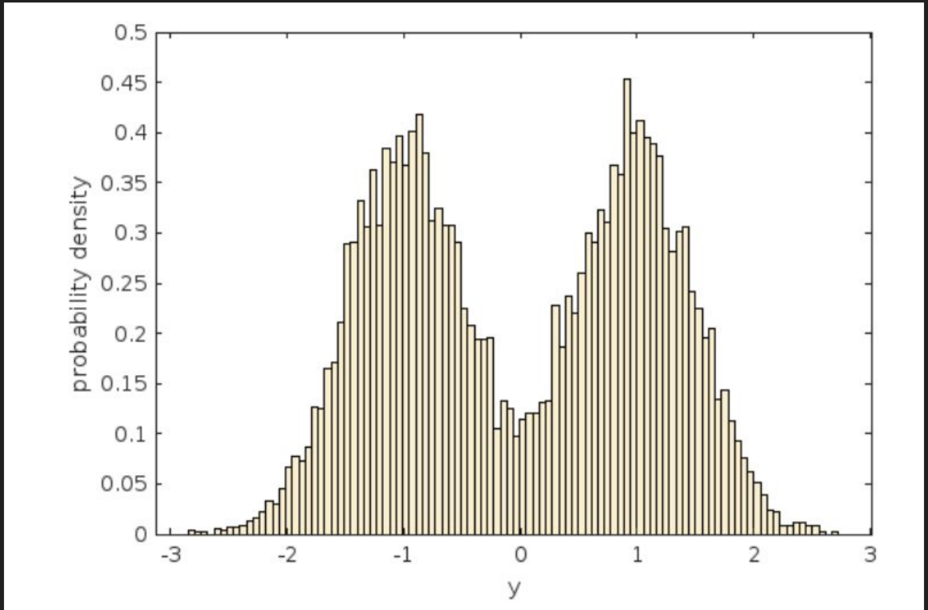

(a) $h(x) = 3x$. (b) $h(x) = x^3$.

Figure 3.3: Simulation for **d)**. In this simulation, $x \sim \mathcal{N}(-5, 1)$ and $r \sim \mathcal{N}(0, 1)$. When $h(x) = 3x$, $y|x \sim \mathcal{N}(-15, 10)$; when $h(x) = x^3$, we cannot describe $p(y)$. In this case, whether we can describe $p(y)$ depends on whether $h(x)$ is a linear or nonlinear function. As for $p(y|x)$, the simulation results will be the same as Figure 3.2 if $x = 2$.

# MMSE and MAP estimators

a) Gaussian mixture distribution

$$y \sim 0.5 \cdot \mathcal{N}(-1, 0.5^2) + 0.5 \cdot \mathcal{N}(1, 0.5^2)$$



# MMSE and MAP estimators

c)

$$
\begin{aligned}
p(y) &= \int p(y|\theta)p(\theta)\, d\theta \\
&= \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\theta)^2}{2\sigma^2}} \cdot 0.5(\delta(\theta-1)+\delta(\theta+1))\, d\theta \\
&= \int 0.5\delta(\theta-1)\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\theta)^2}{2\sigma^2}} + 0.5\delta(\theta+1)\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\theta)^2}{2\sigma^2}}\, d\theta \\
&= 0.5\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}} + 0.5\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}}
\end{aligned}
$$

# MMSE and MAP estimators

d)

$$p(\theta|y) = \frac{p(y|\theta)\,p(\theta)}{p(y)}$$

$$\begin{aligned}
p(\theta = 1|y) &= \frac{e^{\frac{-(y-1)^2}{2\sigma^2}}}{e^{\frac{-(y-1)^2}{2\sigma^2}} + e^{\frac{-(y+1)^2}{2\sigma^2}}} \\
&= \frac{e^{\frac{-y^2}{2\sigma^2}}\, e^{\frac{2y}{2\sigma^2}}\, e^{\frac{-1}{2\sigma^2}}}{e^{\frac{-y^2}{2\sigma^2}}\, e^{\frac{2y}{2\sigma^2}}\, e^{\frac{-1}{2\sigma^2}} + e^{\frac{-y^2}{2\sigma^2}}\, e^{\frac{-2y}{2\sigma^2}}\, e^{\frac{-1}{2\sigma^2}}} \\
&= \frac{e^{\frac{-y^2}{2\sigma^2}}\, e^{\frac{2y}{2\sigma^2}}\, e^{\frac{-1}{2\sigma^2}}}{e^{\frac{-y^2}{2\sigma^2}}\, e^{\frac{-1}{2\sigma^2}} \left( e^{\frac{2y}{2\sigma^2}} + e^{\frac{-2y}{2\sigma^2}} \right)} \\
&= \frac{e^{\frac{y}{\sigma^2}}}{e^{\frac{y}{\sigma^2}} + e^{\frac{-y}{\sigma^2}}}
\end{aligned}$$

$$p(\theta = -1|y) = \frac{e^{\frac{-y}{\sigma^2}}}{e^{\frac{y}{\sigma^2}} + e^{\frac{-y}{\sigma^2}}}$$

# MMSE and MAP estimators

e) MMSE estimator is given as the tanh function with argument dependent on y and $\sigma$

$$\begin{aligned}\hat{\theta}_{MMSE} &= p(\theta = 1|y) - p(\theta = -1|y) \\ &= \frac{e^{\frac{y}{\sigma^2}} - e^{\frac{-y}{\sigma^2}}}{e^{\frac{y}{\sigma^2}} + e^{\frac{-y}{\sigma^2}}} \\ &= \tanh\left(\frac{y}{\sigma^2}\right)\end{aligned}$$

# MMSE and MAP estimators

f) MAP estimator

$$\hat{\theta}_{MAP} = arg\max_{\theta} p(\theta|y) = \begin{cases} 1 & \text{if } \; p(\theta = 1|y) > p(\theta = -1|y) \\ -1 & \text{if } \; p(\theta = 1|y) < p(\theta = -1|y) \end{cases}$$

$$\hat{\theta}_{MAP} = \begin{cases} 1 & \text{if } y > 0 \\ -1 & \text{if } y < 0 \end{cases}$$

g) Given $\sigma^2 = 0.5^2$ and $y = 0.7$, $\hat{\theta}_{MAP} = 1$ since $y > 0$.

$\hat{\theta}_{\text{MMSE}} = \tanh\left(\frac{0.7}{0.25}\right) \approx 0.9926$ → continuous