

Week-7 Questions

Prof . Prathosh AP and Chandan J

July 2025

1. Latent Dimension in Diffusion Models will be

- (A) Less than that of Data Dimension
- (B) More than that of Data Dimension
- (C) Equal to Data Dimension

Answer: (C)

2. Choose the Correct answer

- (A) Parameters of Forward and Reverse diffusion process are learnt
- (B) Parameters of Forward and Reverse diffusion process are not learnable
- (C) Parameters of Forward is not learnable and Reverse diffusion process is learnt.
- (D) Parameters of Forward process is learnt and Reverse diffusion process is not learnable.

Answer: (C)

3. Which term in the DDPM ELBO formulation is independent of model parameters

- (A) Reconstruction Term is independent of parameters
- (B) Prior matching term is independent of parameters
- (C) Denoising Matching Term is independent of parameters

Answer: (B)

4. When $T=1$ diffusion models are same as

- (A) Vanilla VAE
- (B) Beta(10) VAE
- (C) Info VAE
- (D) VQ VAE

Answer: (A)

5. In a Variational Autoencoder, the Evidence Lower Bound (ELBO) is used as a surrogate for the marginal likelihood. Which of the following statements about ELBO is correct?
- (A) The ELBO is maximized directly, which is equivalent to minimizing the marginal likelihood.
 - (B) Maximizing the ELBO implicitly minimizes the Jensen-Shannon divergence between the approximate posterior and prior.
 - (C) The ELBO maximizes the sum of the reconstruction loss and a regularization term, which encourages structured latent spaces.
 - (D) The ELBO approximates the Wasserstein distance between $q(z|x)$ and $p(z|x)$

Answer: (C)

6. In a VAE with a Gaussian latent prior, the decoder is often implemented as a neural network that outputs the mean and variance of a Gaussian distribution over the observation space. Why is this choice beneficial?
- (A) It enforces that the decoder output is discrete, facilitating easier sampling from the latent space.
 - (B) This allows the model to capture uncertainty in the data, as each reconstruction is sampled from a distribution.
 - (C) The Gaussian output improves the decoder's ability to minimize the total variation distance.
 - (D) The ELBO approximates the Wasserstein distance between $q(z|x)$ and $p(z|x)$

Answer: (B)

7. Which of the following statements accurately describes the role of the KL divergence term in the VAE objective function?
- (A) The KL divergence term controls the smoothness of the reconstructed data in the observation space.
 - (B) It minimizes the distance between the empirical data distribution and the reconstructed distribution.
 - (C) It regularizes the encoder to align the approximate posterior $q(z|x)$ with prior $p(z)$ encouraging latent space structure.
 - (D) It ensures that the latent variables z , are independent, reducing the variance in the output.

Answer: (C)

8. In the VAE framework, the posterior distribution $q(z|x)$ is often modeled as a Gaussian distribution. Why is this choice advantageous for VAEs?

- (A) It allows the KL divergence term in the loss function to be computed in closed form.
- (B) It ensures that the latent space is discrete, making it easier to interpret.
- (C) It guarantees that the model will always generate high-quality samples.
- (D) It simplifies the reconstruction term in the loss function, allowing it to be computed analytically.

Answer: (A)

9. Which of the following modifications in a VAE would likely help to produce disentangled representations in the latent space?
- (A) Increasing the weight of the reconstruction loss term to ensure more accurate reconstructions.
 - (B) Introducing a hierarchical prior in the latent space to enforce multi-scale dependencies among latent variables.
 - (C) Adding a total correlation regularization term to the loss function to encourage independence among latent dimensions.
 - (D) Using a non-Gaussian likelihood function to encourage more diverse samples.

Answer: (C)

10. Beta > 1 in Beta VAE does
- (A) applies a stronger constraint on the latent bottleneck and limits the representation capacity of Z .
 - (B) applies a weaker constraint on the latent bottleneck and limits the representation capacity of Z .
 - (C) applies a weaker constraint on the latent bottleneck and does not limits the representation capacity of Z .
 - (D) applies a stronger constraint on the latent bottleneck and does not limits the representation capacity of Z .

Answer: (A)