1 / 1
points

1.

Select correct statements

☐ One can obtain sample from gaussian distribution using one sample from uniform

Un-selected is correct

☐ One can obtain n multivariate gaussian samples $x \in \mathbb{R}^n$ from n standard 1d gaussian samples.

Un-selected is correct

☐ One can obtain multivariate gaussian sample $x \in \mathbb{R}^n$ from n standard 1d gaussian samples

Correct
Correct: any multivariate gaussian $X \sim \mathcal{N}(\mu, \Sigma)$ can be represented as $\mu + \Sigma^{1/2}\epsilon$, where $\epsilon \sim \mathcal{N}(0, \mathrm{I})$.

☐ One can obtain sample from exponential distribution using one sample from uniform

Correct
Correct: If $X \sim \mathcal{F}$ is a random variable with CDF $F(x)$ that can be inverted analytically then $F^{-1}(u) \sim \mathcal{F}$ where $u \sim \mathrm{U}[0, 1]$. Correspondingly $-\frac{\log u}{\lambda} \sim Exp(\lambda)$.

1 / 1
points

2.

What is a time complexity of an algorithm for sampling a random number from an arbitrary discrete distribution with support $\{1, \ldots, N\}$

O(N)

Correct
Correct: prior to sampling we need to compute cumulative sums which is O(N) operations.

O(log N)

O(N log N)

O(1)

---

1 / 1
points

3.

What we can use Monte-Carlo method for?

Compute an integral of an arbitrary function over a simple area (e.g. a multidimensonal cube)

Correct
Correct: just sample random points uniformly in the area and average the function values in those points. Note that usually this is not the best approach, although very scalable.

Estimate the expected values of arbitrary random variables

Correct
Correct: this is what Monte-Carlo method is for.

Do full bayesian inference to estimate the uncertainty of you model.

Correct

Correct: see example in the lecture.

Compute the exact mode of a posterior distribution (MAP-estimation)

Un-selected is correct

1 / 1 points

4.

Which of the statements below are correct?

Any sequence of random variables $X_n : n \in \mathbb{N}$ can be considered as a Markov chain.

Un-selected is correct

Any Markov chain converges to a stationary distribution

Un-selected is correct

Any Markov chain is a sequence of discrete random variables, for example: $\{0, 1, 0, 0, 1, 0, 0, \ldots\}$.

Un-selected is correct

Markov chain does not "remember" states other than current

Correct

Correct.

- [ ] All elements X_n of a Markov chain $X_n : n \in \mathbb{N}$ are independent random variables.

Un-selected is correct

---

1 / 1
points

5.

Which of the statements below are correct?

- [ ] MCMC techniques are used when ones cannot perform bayesian inference analytically

Correct
Correct.

- [ ] MCMC provides i.i.d. samples from desired distribution

Un-selected is correct

- [ ] MCMC can be used to sample from the distribution known up to a normalization constant

Correct
Correct.

---

1 / 1
points

6.

Which of the statements below are correct?

- [ ] Gibbs sampling is a special case of a Metropolis-Hastings algorithm.

Correct
Correct. Gibbs sampling is a special case of MH with acceptance rate equal to 1.

- [ ] Gibbs sampling reduces multidimensional sampling to one-dimensional sampling.

Correct
Correct.

- [ ] Each iteration of Gibbs sampling changes only one coordinate of a latent vector

Correct
Correct.

- [ ] Gibbs sampling converges really fast because it provides very uncorrelated samples compared to Metropolis-Hastings algorithm

Un-selected is correct

1 / 1
points

7.

Which of the following is random in Bayesian Neural Networks?

- [ ] Weights of the network w

Correct
Correct

- [ ] Prediction of the network y given fixed input x

Correct
Correct. Prediction of the network y depends on the weights which are random variables.

- [ ] Number of units on each layer of the network

Un-selected is correct

- [ ] Number of active layers of the network

Un-selected is correct

---

1 / 1
points

8.

What is a good way to train (find the posterior distribution $p(w|D)$) Bayesian Neural Network?

- ( ) Iteratively sample each weight from the conditional distribution given all other weights and the data.
- ( ) Compute the posterior distribution $p(w|D)$ analytically.
- ( ) Run the stochastic gradient descent perturbing all network weights with independent Gaussian noise after each iteration.

Correct
Correct. This algorithm is called Langevin Monte Carlo and is proved to converge to the true posterior.

---

1 / 1
points

9.

What does the word “Collapsed” means in the Collapsed Gibbs Sampling algorithm?

- It means that posterior distribution over some of the variables is computed analytically, while other variables are sampled using Gibbs Sampling.

Correct
Correct

- It means that the posterior approximation *collapses* to the are posterior distribution.
- It means that we train the model on the subsample of the original data.

---

1 / 1
points

10.

Which of the variables are randomly sampled in Collapsed Gibbs Sampling for LDA?

- Z

Correct

Correct. Everything else can be computed analytically.

- Φ, Θ
- Φ, Θ, Z
- Φ