(1) Introduction

1. Give a definition of Data Science.
2. Give a definition of Statistics.
3. Name three central postulates of Probability theory.
4. Name three central postulates of Frequentist inference.
5. Name three scientists involved in the development of Frequentist statistics.
6. Name three central postulates of Bayesian inference.
7. Name three scientists involved in the development of Bayesian statistics.
8. Name five typical topics in Statistics.
9. Name three topics commonly discussed in Machine Learning.
10. Name three topics commonly discussed in Artificial Intelligence.

(2) Probability spaces

1. Write down the definition of a probability space.
2. Give two interpretations for the probability $\mathbb{P}(A)$ of an event A.
3. Sketch the probability space model of throwing a dice.
4. Write down the definition of a probability measure.
5. For a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $A, B \in \mathcal{A}$. What is the probability of the event that A or B are true?
6. Write down the definition of the independence of two events A and B and the definition of the independence of a set of events $\{A_i | i \in I\}$ with index set I.
7. Write down the definition of the conditional probability of an event given another B.
8. State the law of total probability.
9. What is the conditional probability of an event given an event B, if A and B are independent events? Justify your answer.
10. Write down and prove Bayes theorem.

(3) Random variables

1. Write down the definition of a random variable.
2. Write down the definition of a discrete random variable and a probability mass function (PMF).
3. Write down the definition of a continuous random variable and a probability density function (PDF).
4. Write down the definition of the cumulative distribution function (CDF) of a random variable.
5. Express the value $P(x)$ of the CDF of a discrete random variable X in terms of its PMF.
6. Express the value $P(x)$ of the CDF of a continuous random variable X in terms of its PDF.
7. Express the value $p(x)$ of the PDF of a continuous random variable X in terms of its CDF.
8. Write down the PDF and CDF of a Gaussian random variable
9. State three properties of CDFs.
10. Write down the definitions of the inverse CDF and the quantile functions.

(4) Joint distributions

1. Write down the definition of a joint PMF of two discrete random variables X and Y with finite outcome set.
2. Write down the definition of the joint PDF of two continuous random variables X and Y each taking values in $\mathbb{R}$.
3. Write down the definitions of the marginal PMFs and PDFs of a joint distribution of two random variables X and Y with joint PMF or PDF $p_{X,Y}$.
4. Write down the definition of the independence of two random variables X and Y.
5. Write down the necessary and sufficient condition for the independence of two random variables X and Y with joint PMF/PDF $p_{X,Y}$.
6. Write down the definitions of conditional PMFs and PDFs.
7. Write down the definition of a multivariate Gaussian PDF and comment on the meaning of its parameters.
8. Write down the definition of the independence of n random variables X_i.
9. What does it mean for n random variables $X_1, ..., X_n$ to be i.i.d.?
10. What does it mean for $X_1, ..., X_n$ to be a random sample of size n from p?

(5) Transformations

1. Let X be a discrete random variable and $Y = f(X)$ a transformation of X. Express the PMF p_Y of Y in terms of the distribution of X.
2. Let X be a continuous random variable with PDF p_X and let $Y = f(X)$ be a transformation of X. Write down the direct calculation procedure to derive the PDF p_Y of Y.
3. Write down the probability integral transform theorem.
4. How can a uniform random number generator be used to create random numbers with arbitrary distribution?
5. Write down the univariate probability density function transform theorem.
6. Write down the univariate probability density function transform theorem for linear functions.
7. Write down the Z-transformation of a univariate Gaussian random variable.
8. Write down the multivariate probability density function transform theorem.
9. Write down the multivariate probability density function transform theorem for linear functions.
10. Let an n-dimensional random vector X be distributed according to a multivariate Gaussian distribution, $X \sim N(\mu_x, \Sigma_x)$. Let $A \in \mathbb{R}^{n \times n}$ be a matrix of full column-rank. How is the random vector $Y := AX$ distributed?

(6) Expectation and covariance

1. Discuss the intuition of the expected value of a random variable.
2. What does it mean for the expectation of a random variable to exist?
3. Compute the expectation of a Bernoulli random variable.
4. State the linearity and multiplication properties of expectations.
5. Write down $\mathbb{E}(X^2)$ in terms of the variance and expectation of the random variable X.
6. For constant a, what is $\mathbb{V}(aX)$?
7. Write down the definition of the covariance and correlation of two random variables X and Y.
8. Express the covariance of two random variables X and Y in terms of expectations.
9. What is the variance of the sum of two random variables X and Y, if X and Y are independent and in general?
10. What is the variance of the difference of two random variables X and Y, if X and Y are independent and in general?

(7) Inequalities and limits

1. Write down the Markov inequality.
2. Write down the Chebychev inequality.
3. Write down Jensen's inequality for concave functions.
4. Write down the Cauchy-Schwarz inequality.
5. Write down the Correlation inequality.
6. Write down the definition of convergence in probability.
7. Write down the definition of almost-sure convergence.
8. Write down the definition of convergence in distribution.
9. Write down the Weak Law of Large Numbers.
10. Write down the Strong Law of Large Numbers.
11. Write down the Lindenberg-Lévy form of the Central Limit Theorem.
12. Write down the Liapunov form of the Central Limit Theorem.

(8) Maximum likelihood estimation

1. Define the notion of a point estimator.
2. Write down the definitions of the likelihood and log likelihood functions.
3. Write down the definition of the maximum likelihood estimator.
4. Write down the general maximum likelihood procedure for parametric statistical models.
5. Derive the maximum likelihood estimator for the parameter of a Bernoulli distribution.
6. Derive the maximum likelihood estimator for the parameters of a Gaussian distribution.
7. Formulate the univariate Newton-Raphson method.
8. Formulate the multivariate Newton-Raphson method.
9. Write down the Fisher scoring algorithm.

(9) Finite estimator properties

1. Define the bias of an estimator. When is an estimator unbiased?
2. Write down the definition of the variance of an estimator.
3. Define the standard error of an estimator.
4. Define the standard error of the mean.
5. Write down the definitions of the score function of a random variable/sample.
6. Write down the definitions of the Fisher information of a random variable/sample.
7. Write down the definitions of the expected Fisher information of a random variable/sample.
8. What are the expected value and the variance of the score function of a random variable?
9. Formulate the Cramér-Rao bound theorem.
10. Write down the bias-variance decomposition for the mean squared error of an estimator.

(10) Asymptotic estimator properties

1. Write down the definition of an asymptotically unbiased estimator.
2. Write down the definition of a consistent estimator.
3. State the mean squared error criterion for estimator consistency.
4. State the bias and variance criterion for estimator consistency.
5. Write down the definition of an asymptotically normally distributed estimator.
6. Write down the definition of an asymptotically efficient estimator.
7. Name five properties of maximum likelihood estimators.
8. Given an example of a biased maximum likelihood estimator.
9. Sketch the proof of the consistency of maximum likelihood estimators
10. Sketch the proof of the asymptotic efficiency of maximum likelihood estimators.

(11) Confidence intervals

1. Write down the definition of the T statistic and state its distribution.
2. Write down the definition of the Wald statistics and state their distribution.
3. Define the δ-confidence interval.
4. Give two interpretations of δ-confidence intervals.
5. Define the notions of exact and approximate pivots and δ-confidence intervals.
6. State the steps involved in the typical construction of confidence intervals.
7. Write down the formula of the 95%-confidence interval for the expectation parameter of a univariate Gaussian distribution with known variance.
8. Write down the formula of the 95%-confidence interval for the expectation parameter of a univariate Gaussian distribution with unknown variance.
9. Write down the formula of the 95%-confidence interval for the variance parameter of a univariate Gaussian distribution.
10. Write down the formula of an approximate 95%-confidence interval for a parameter based on a maximum likelihood estimator.

(12) Hypothesis testing

1. Define the notions of test hypotheses, as well as simple, composite, and nil hypotheses.
2. Define the notion of a statistical test and of a standard statistical test.
3. Write down the definition of a one-sided critical value-based test.
4. Write down the definition of a two-sided critical value-based test.
5. Define the notions of Type I and Type II test errors.
6. Define the size, specificity, power, and significance level of a test.
7. Define the notions of a conservative, exact, and liberal test.
8. Write down the definition of the test quality and power function.
9. State the typical procedure for constructing a hypothesis test.
10. Formulate the duality of confidence intervals and hypotheses tests.

(13) Conjugate inference

1. Write down the definition of a probabilistic model, a generative model, a prior distribution, a likelihood, and a posterior distribution.
2. Write down the distribution of n conditionally independent and identically distributed random variables $X_i, i = 1, ..., n$ given a parameter random variable θ.
3. Describe the differences and similarities between batch and recursive Bayesian estimation.
4. Write down the definition of the marginal data likelihood (model evidence).
5. Given two probabilistic models and a set of data observations, write down the Bayes factor.
6. Write down the definition of the posterior predictive distribution.
7. Write down the definition of a loss function, the expected posterior loss, and a Bayes estimator.
8. Write down the Bayes estimator under a quadratic loss function.
9. Write down the Bayes estimator under zero-one loss function.
10. Write down the definition of a conjugate family of distributions.

(14) Numerical methods

1. Name two quantities in Bayesian inference that often necessitate numerical integration.
2. Name an example for a quadrature rule.
3. What is the difference between the right rule and the midpoint rule in Riemann sum-based numerical integration?
4. State Laplace's integral approximation method.
5. Write down the Laplace approximation of a posterior expectation of the form $\mathbb{E}_{p(\theta|x_{1:n})}(f(\theta))$.
6. Write down the definition of the Monte Carlo estimator of an integral $I = \int_{\mathcal{X}} f(x)p(x)\,dx$.
7. State the importance sampling identity.
8. Write down the acceptance-rejection algorithm.

(15) Variational inference

1. Define the variational inference problem.
2. Write down the log model evidence decomposition.
3. Write down the definition of the evidence lower bound.
4. Write down the definition and two properties of the Kullback-Leibler divergence.
5. State the evidence lower bound theorem.
6. Describe two approaches of using the evidence lower bound theorem for solving the variational inference problem.
7. Define the concept of a mean-field approximation in variational inference.
8. State the free-form mean-field variational inference theorem.
9. Write down the general CAVI algorithm.
10. Define the concept of fixed-form variational inference.