These exercises were part of the Statistics for Data Science course 2019/20 held at the FU Berlin. They are intellectual property of Prof. Dirk Ostwald and can be found here.

Probability theory

(1) Probability spaces

1. Consider the probability space model of tossing a fair dice. Let $A = \{2, 4, 6\}$ and $B = \{1, 2, 3, 4\}$ be two events. Then, $\mathbb{P}(A) = 1/2$, $\mathbb{P}(B) = 2/3$ and $\mathbb{P}(A \cap B) = 1/3$. Since $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, the events A and B are independent. Simulate draws from the outcome space and verify that $\hat{\mathbb{P}}(A \cap B) = \hat{\mathbb{P}}(A)\hat{\mathbb{P}}(B)$, where $\hat{\mathbb{P}}(E)$ denotes the proportion of times an event E occurs in the simulation. Next, identify two events A and B that are not independent. Analytically, evaluate $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(A \cap B)$, $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$ and verify these values using the simulation. Document your results.

2. Consider the probability space model of tossing a fair coin twice and the events "heads appears on the first toss", "heads appears on the second toss", and "both tosses have the same outcome". Simulate draws from the outcome space and verify that (1) the probability of each of the events is 1/2, (2) the probability for the co-occurrence of pairs of the events, as well as of all events is 1/4.

(2) Random variables

1. Visualize the probability mass function of a Bernoulli random variable and the probability density function of a Gaussian random variable. On top, visualize histograms of many samples from each random variable for the same parameter settings. For the Gaussian random variable, document how you normalize the histogram.

2. Visualize the probability density functions of a Beta random variable and a Gamma random variable (**?**, Definitions 5.8.2 and 5.7.2). On top, visualize appropriately normalized histograms of many samples from each random variable for the same parameter settings.

(3) Joint distributions

1. Write a simulation that demonstrates that the marginal distributions of a bivariate Gaussian distribution with expectation parameter $\mu = (1, 2)^T$ and covariance matrix parameter $\Sigma = \begin{pmatrix} 0.3 & 0.2 \\ 0.2 & 0.5 \end{pmatrix}$ are given by univariate Gaussian distributions with expectation parameters $\mu_1 = 1, \mu_2 = 2$ and variance parameters $\sigma^2 = 0.3$ and $\sigma^2 = 0.5$, respectively. For the simulation, make use of multivariate Gaussian probability density and random number generators. Visualize and document your results.

2. Write a simulation that verifies that obtaining samples from 5 independent univariate Gaussian distributions with parameters $\mu_i, \sigma_i^1 > 0, i = 1, ..., 5$ is equivalent to obtaining samples from a 5-dimensional multivariate Gaussian distribution with the appropriately specified parameters $\mu \in \mathbb{R}^5$ and $\Lambda \in \mathbb{R}^{n \times n}$. Visualize and document your results.

3. Write a simulation that verifies the analytical results on joint and conditional Gaussian distributions for the case of univariate X and Y, $A := 2$ and $b := 1$ by visually displaying the respective PDFs and their large sample normalized histograms.

(4) Random variable transformations

1. Write a program that generates pseudo-random numbers from an exponential distribution using a uniform pseudo-random number generator and the probability integral transform theorem. Visualize and document your results.
2. Let $X \sim N(0,1)$ and let $Y = \exp(X)$. Evaluate the PDF of Y analytically and verify your evaluation using a simulation based on drawing random numbers from $N(0,1)$. Visualize and document your results.
3. Let $X \sim N(0,1)$ and let $Y = X^2$. By simulation, validate that Y is distributed according to a chi-squared distribution with one degree of freedom. Next, let $X_1, ..., X_{10} \sim N(0,1)$ and let $Y = \sum_{i=1}^{10} X_i^2$. By simulation, validate that Y is distributed according to a chi-squared distribution with ten degrees of freedom.

(5) Expectation, (co)variance, inequalities, limits

1. Write a simulation that validates the Weak Law of Large Numbers. Visualize and document your results.
2. Write a simulation that validates the Strong Law of Large Numbers. Visualize and document your results.
3. Write a simulation that validates the Central Limit Theorem in the Lindenberg-Lévy form. Visualize and document your results.
4. Write a simulation that validates the Central Limit Theorem in the Liapunov form. Visualize and document your results.

Frequentist inference

(6) Foundations and maximum likelihood

1. Write a program that implements a Fisher scoring algorithm for the maximum likelihood estimation of the parameters of a simple linear regression model. Compare the results with the analytical estimation of the parameters. Visualize and document your results.
2. Let $X_1, ..., X_n \sim \text{Bern}(\mu)$ be $n = 20$ i.i.d. Bernoulli random variables. Using an optimization routine of your choice, formulate and implement the numerical maximum likelihood estimation of μ for true, but unknown values of $\mu = 0.7$ and $\mu = 1$ based on $X_1, ..., X_n$.
3. Let $X_1, ..., X_n \sim \text{Bern}(\mu)$. For a large number n, sample the $X_1, ..., X_n$ and evaluate the maximum likelihood estimator $\hat{\mu}^{ML}$. Repeat this m times and create a histogram of the realized $\hat{\mu}_1^{ML}, ..., \hat{\mu}_m^{ML}$. Visualize and document your results.

(7) Finite-sample estimator properties

1. For $X_1, ..., X_n \sim N(\mu, \sigma^2)$ implement a simulation which validates the unbiasedness of the sample mean, the unbiasedness of the sample variance, the biasedness of the sample standard deviation, and the biasedness of the maximum likelihood variance parameter estimator.
2. For $X_1, ..., X_n \sim B(\mu)$ implement a simulation which validates the unbiasedness of the sample mean, the unbiasedness of the sample variance, the biasedness of the sample standard deviation, and the biasedness of the maximum likelihood variance parameter estimator.
3. Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$ implement a simulation that validates the bias-variance decompositions of the mean squared errors of the maximum likelihood estimator of σ^2, the sample variance S^2, and the estimator $\hat{\sigma}^2$ introduced in the first theoretical exercise above (**?**, Example 8.7.6).

(8) Asymptotic estimator properties

1. Write a simulation that verifies the asymptotic unbiasedness of the maximum likelihood estimator for the variance parameter of a univariate Gaussian distribution. Include a verification of the unbiasedness of the sample variance.
2. Write a simulation that verifies the asymptotic efficiency of the maximum likelihood estimator for the parameter of a Bernoulli distribution.
3. Write a simulation that verifies the asymptotic efficiency of the maximum likelihood estimator for the variance parameter of a univariate Gaussian distribution.

(9) Confidence intervals

1. Write a simulation that verifies that the T statistic is distributed according to a t-distribution with $n-1$ degrees of freedom.
2. Write a simulation that verifies that the 95%-confidence interval for the expectation parameter of a Gaussian distribution with unknown variance comprises the true, but unknown, expectation parameter in $\approx 95\%$ of its realizations.
3. Write a simulation that verifies that the approximate 95%-confidence interval for the expectation parameter of a Bernoulli distribution comprises the true, but unknown, expectation parameter in $\approx 95\%$ of its realizations.

(10) Hypothesis testing

1. By means of simulation, show that a T test of significance level α' is an exact test.
2. By means of simulation, validate the power function of the T test.
3. By means of simulation, demonstrate that the δ-confidence interval-based test for the expectation parameter of univariate Gaussian distribution is of significance level $\alpha' = 1 - \delta$.

(11) Nonparametric inference

1. Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$. Implement a histogram density estimator for the PDF of $X_i := X_1$. Visualize the relationship between the sample size n and the binwidth b in determining the quality of the estimate.
2. Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$. Implement a kernel density estimator for the PDF of $X_i := X_1$. Visualize the relationship between the sample size n and the bandwidth b in determining the quality of the estimate.
3. For $X_1, ..., X_n \sim N(\mu, \sigma^2)$ use the bootstrap to estimate the variance of the T-statistic

$$T = \sqrt{n}\frac{\bar{X}_n - \mu}{S_n}. \tag{1}$$

   Compare the bootstrap estimates to the analytical variance $\mathbb{V}(T) = \frac{n}{n-2}$ for varying $n > 2$ and varying number of bootstrap samples.

Bayesian inference

(12) Foundations and conjugate inference

1. For $n = 10$, implement batch and recursive Bayesian estimation for the Beta-Binomial model. Compare the results based on identical samples.

2. Using simulation, study the bias and consistency properties of the posterior expected value of the Beta-Binomial model.

3. Using simulation, study the bias and consistency properties of the posterior expected value of the Gaussian-Gaussian model.

(13) Bayesian filtering

1. Evaluate the sum of squared deviations of the filtered distribution expectation parameters from the true, but unknown, state values as a function of the transition and observation variance parameters, respectively, for the latent Gaussian walk model using the exemplary code provided with the lecture.

2. Using the exemplary code provided with the lecture, evaluate the variance parameters of the filtered and smoothed distributions as functions of the transition and observation variance parameters, respectively, for the latent Gaussian walk model using the exemplary code provided with the lecture.

(14) Numerical methods

1. Estimate the expected value of a Beta(α, β) for varying values of α and β by means of Monte Carlo integration by using a Beta distribution random number generator. Compare the results to the true expected values.

2. Estimate the expected value of a Beta(α, β) for varying values of α and β by means of Monte Carlo integration using an importance sampling scheme and a uniform random number generator.

3. Use an acceptance-rejection algorithm to sample random numbers from Beta$(5, 1)$.

(15) Variational inference

1. Evaluate and visualize the KL divergence between two Gaussian PDFs as well as between two Gamma PDFs for varying parameter settings of the distributions. Closed form solutions for these KL divergence are available from the literature.

2. Implement the free-form CAVI algorithm for the Gaussian-Gamma model as introduced in the lecture. Visualize the likelihood, prior, approximate posterior, iterative variational distributions, as well as the ELBO.

3. Implement fixed-form algorithm for the nonlinear Gaussian model as introduced in the lecture. Visualize the likelihood, prior, approximate posterior, iterative variational distributions, as well as the ELBO approximation function.