Statistical Methods in Particle Physics

Quiz on chapter 1: Basics

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

Probabilities satisfy $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint according to the

1. Kolmogorov axioms
2. Chebyshev axioms
3. Markov axioms
4. Poincare axioms

Probabilities satisfy $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint according to the

1. Kolmogorov axioms
2. Chebyshev axioms
3. Markov axioms
4. Poincare axioms

The mode of a distribution is

1. the value separating the higher half of the distribution from the lower half
2. the expectation value of the distribution
3. the value x at which f(x) takes its maximum value
4. the standard deviation of the distribution

The mode of a distribution is

1. the value separating the higher half of the distribution from the lower half
2. the expectation value of the distribution
3. the value x at which f(x) takes its maximum value
4. the standard deviation of the distribution

The variance of a continuous distribution can be calculated as

1. $V(x) = \langle x^2 \rangle + \langle x \rangle^2$
2. $V(x) = \langle x^2 \rangle - \langle x \rangle^2$
3. $V(x) = \langle x^2 + \langle x \rangle^2 \rangle^2$

$$V(x) = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

The variance of a continuous distribution can be calculated as

1. $V(x) = \langle x^2 \rangle + \langle x \rangle^2$
2. $V(x) = \langle x^2 \rangle - \langle x \rangle^2$
3. $V(x) = \langle x^2 + \langle x \rangle^2 \rangle^2$

$$V(x) = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

The covariance of two random variables x and y is defined as

1. $\mathrm{cov}[x, y] = E[(x - \langle x \rangle)^2 (y + \langle y \rangle)^2]$
2. $\mathrm{cov}[x, y] = E[(x - \langle x \rangle)^2 (y - \langle y \rangle)^2]$
3. $\mathrm{cov}[x, y] = E[(x - \langle x \rangle)(y - \langle y \rangle)]$

$$\mathrm{cov}[x, y] = E[(x - \langle x \rangle)^2] \cdot E[(y - \langle y \rangle)^2]$$

The covariance of two random variables x and y is defined as

1. $\mathrm{cov}[x, y] = E[(x - \langle x \rangle)^2 (y + \langle y \rangle)^2]$

2. $\mathrm{cov}[x, y] = E[(x - \langle x \rangle)^2 (y - \langle y \rangle)^2]$

3. $\mathrm{cov}[x, y] = E[(x - \langle x \rangle)(y - \langle y \rangle)]$

$\mathrm{cov}[x, y] = E[(x - \langle x \rangle)^2] \cdot E[(y - \langle y \rangle)^2]$

Statistical Methods in Particle Physics

Quiz on chapter 2: Probability Distributions

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

The $\pm1\sigma$ interval around the mean of a Gaussian corresponds to a probability of about

1. 32%
2. 36%
3. 68%
4. 95%

The ±1σ interval around the mean of a Gaussian corresponds to a probability of about

1. 32%
2. 36%
3. 68%
4. 95%

The Central Limit Theorem states that

1. histograms approach the underlying PDF for $n \to \infty$
2. that n! can be calculated as $\Gamma(n+1)$
3. a binomial distributions can be approximated by Poisson distribution under certain conditions
4. the sum of n random variables approaches a Gaussian distribution for $n \to \infty$

The Central Limit Theorem states that

1. histograms approach the underlying PDF for $n \to \infty$
2. that $n!$ can be calculated as $\Gamma(n+1)$
3. a binomial distributions can be approximated by Poisson distribution under certain conditions
4. the sum of n random variables approaches a Gaussian distribution for $n \to \infty$

The energy loss of a charged particle in a thin material layer can be described by the

1. exponential distribution
2. Lorentz distribution
3. logarithmic distribution
4. Landau distribution

The energy loss of a charged particle in a thin material layer can be described by the

1. exponential distribution
2. Lorentz distribution
3. logarithmic distribution
4. Landau distribution

In particle physics, the mass distribution of a resonance can be described by a

1. exponential distribution
2. Lorentz distribution
3. negative binomial distribution
4. χ^2 distribution

In particle physics, the mass distribution of a resonance can be described by a

1. exponential distribution
2. Lorentz distribution
3. negative binomial distribution
4. χ^2 distribution

The difference of two Gaussian distributed random variables follows a

1. Student's t distribution
2. Cauchy distribution
3. Gaussian distribution
4. none of the above

The difference of two Gaussian distributed random variables follows a

1. Student's t distribution
2. Cauchy distribution
3. Gaussian distribution
4. none of the above

The product of two Gaussian distributed random variables follows a

1. Student's t distribution
2. χ^2 distribution
3. Gaussian distribution
4. none of the above

The product of two Gaussian distributed random variables follows a

1. Student's t distribution
2. χ^2 distribution
3. Gaussian distribution
4. none of the above

The expectation value of a χ^2 distribution with n degrees of freedom is

1. n
2. $n\,(n-1)\,/2$
3. $n!$
4. n^2

The expectation value of a χ^2 distribution with n degrees of freedom is

1. n
2. $n\,(n-1)\,/2$
3. $n!$
4. n^2

Statistical Methods in Particle Physics

Quiz on chapter 3: Uncertainties

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

Suppose the average number of proton-proton collisions per bunch crossing at an interaction point of the LHC is 25. What is the variance of the number or collisions per bunch crossing?

1. 5
2. 12.5
3. 25
4. 625

Suppose the average number of proton-proton collisions per bunch crossing at an interaction point of the LHC is 25. What is the variance of the number or collisions per bunch crossing?

1. 5
2. 12.5
3. 25
4. 625

The uncertainty of the sum $z = x + y$ of two uncorrelated variables x and y is given by the square root of

1. the uncertainties of x and y added in quadrature
2. the relative uncertainties of x and y added in quadrature
3. uncertainties of x and y added linearly
4. the absolute values of the relative uncertainties of x and y added linearly

The uncertainty of the sum $z = x + y$ of two uncorrelated variables x and y is given by the square root of

1. the uncertainties of x and y added in quadrature
2. the relative uncertainties of x and y added in quadrature
3. uncertainties of x and y added linearly
4. the absolute values of the relative uncertainties of x and y added linearly

Using linear error propagation, the relative uncertainty of the product $z = x \times y$ of two uncorrelated variables x and y is given by the square root of

1. the uncertainties of x and y added in quadrature
2. the relative uncertainties of x and y added in quadrature
3. the product of the relative uncertainties of x and y
4. the absolute value of the product of the relative uncertainties of x and y

Using linear error propagation, the relative uncertainty of the product $z = x \times y$ of two uncorrelated variables x and y is given by the square root of

1. the uncertainties of x and y added in quadrature
2. the relative uncertainties of x and y added in quadrature
3. the product of the relative uncertainties of x and y
4. the absolute value of the product of the relative uncertainties of x and y

Linear error propagation of two correlated measurements x and y

1. is identical to the case of uncorrelated measurements
2. is possible if the covariance matrix of x and y is known
3. can only be done numerically (Monte Carlo error propagation)
4. is not possible

Linear error propagation of two correlated measurements x and y

1. is identical to the case of uncorrelated measurements
2. is possible if the covariance matrix of x and y is known
3. can only be done numerically (Monte Carlo error propagation)
4. is not possible

In the weighted average of n measurements with uncertainties σ_i, the weight of each measurement is proportional to

1. σ_i
2. σ_i^2
3. $1 / \sigma_i$
4. $1 / \sigma_i^2$

In the weighted average of n measurements with uncertainties σ_i, the weight of each measurement is proportional to

1. σ_i
2. σ_i^2
3. $1 / \sigma_i$
4. $1 / \sigma_i^2$

Statistical Methods in Particle Physics

Quiz on chapter 4: Monte Carlo Methods

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

The name "Monte Carlo method" refers to

1. the inventor Carlo Montego
2. a conference which took place in Monte Carlo
3. the Monte Carlo Casino in Monaco
4. to the formula one race in Monte Carlo

Random numbers generated on a Computer are called pseudo-random numbers because

1. of the limited precision of floating point numbers
2. they result from deterministic algorithms
3. they are only generated in the interval [0,1]
4. they are taken from big look-up tables obtained from throwing real dice

In the inverse transform method to get random numbers from a distribution f(x) one needs to calculate the inverse of

1. f(x)
2. 1/f(x)
3. the first derivative of f(x)
4. the CDF of f(x)

Let r be a random variable uniformly distributed in [0, 1]. To draw random numbers from the PDF $f(x) = 2x$ one can transform r as

1. $\sqrt{r}$
2. r^2
3. $\ln r$
4. r^4

To obtain random points uniformly distributed on the surface of a sphere one needs to uniformly distribute

1. ϕ and θ
2. $\sin\phi$ and θ
3. ϕ and $\cos\theta$
4. ϕ^2 and θ

Monte Carlo integration outperforms other numerical methods in case of

1. multi-dimensional integrals
2. Gaussian integrals
3. positive integrands
4. periodic integrands

Statistical Methods in Particle Physics

Quiz on chapter 5: Parameter Estimation

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

An estimator is biased if

1. the number of data points is finite

2. its expectation value differs from the true value

3. it is a maximum likelihood estimator

4. it has a large variance

An estimator is biased if

1. the number of data points is finite
2. its expectation value differs from the true value
3. it is a maximum likelihood estimator
4. it has a large variance

To obtain an unbiased estimate of the variance of a data sample one has to divide $\sum_{i=1}^{n}(x_i - \bar{x})^2$ by

1. n
2. n^2
3. $n\,(n - 1)$
4. $n - 1$

To obtain an unbiased estimate of the variance of a data sample one has to divide $\sum_{i=1}^{n}(x_i - \bar{x})^2$ by

1. n
2. n^2
3. $n\,(n-1)$
4. $n-1$

The variance of a chi-squared estimator for one parameter is related to

1. the first derivative of the chi-squared function
2. the second derivative of the chi-squared function
3. the logarithm of the chi-squared function
4. the integral of the chi-squared from the measured value to infinity

The variance of a chi-squared estimator for one parameter is related to

1. the first derivative of the chi-squared function
2. the second derivative of the chi-squared function
3. the logarithm of the chi-squared function
4. the integral of the chi-squared from the measured value to infinity

In the large sample limit the likelihood function *L* approaches a

1. Gaussian
2. parabolic function
3. chi-squared distribution
4. logarithmic function

In the large sample limit the likelihood function *L* approaches a

1. Gaussian
2. parabolic function
3. chi-squared distribution
4. logarithmic function

The variance of a maximum likelihood estimator for one parameter θ as obtained from the minimum variance bound is given by

1. $$V[\hat{\theta}] = -\left.\frac{\partial \ln L}{\partial \theta}\right|_{\theta=\hat{\theta}}$$
2. $$V[\hat{\theta}] = -\left.\frac{\partial^2 \ln L}{\partial^2 \theta}\right|_{\theta=\hat{\theta}}$$
3. $$V[\hat{\theta}] = -\frac{1}{\left.\frac{\partial \ln L}{\partial \theta}\right|_{\theta=\hat{\theta}}}$$
4. $$V[\hat{\theta}] = -\frac{1}{\left.\frac{\partial^2 \ln L}{\partial^2 \theta}\right|_{\theta=\hat{\theta}}}$$

The variance of a maximum likelihood estimator for one parameter θ as obtained from the minimum variance bound is given by

1. $$V[\hat{\theta}] = - \left. \frac{\partial \ln L}{\partial \theta} \right|_{\theta=\hat{\theta}}$$

2. $$V[\hat{\theta}] = - \left. \frac{\partial^2 \ln L}{\partial^2 \theta} \right|_{\theta=\hat{\theta}}$$

3. $$V[\hat{\theta}] = - \frac{1}{\left. \frac{\partial \ln L}{\partial \theta} \right|_{\theta=\hat{\theta}}}$$

4. $$V[\hat{\theta}] = - \frac{1}{\left. \frac{\partial^2 \ln L}{\partial^2 \theta} \right|_{\theta=\hat{\theta}}}$$

Statistical Methods in Particle Physics

Quiz on chapter 6: Hypothesis Testing

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

A simple hypothesis

1. can be formulated analytically
2. has no free parameters
3. is rejected with a probability larger than 68%
4. can be tested with relatively small data samples

A simple hypothesis

1. can be formulated analytically
2. has no free parameters
3. is rejected with a probability larger than 68%
4. can be tested with relatively small data samples

A variable that is a function of the data alone and that can be used to test a hypothesis is called

1. run test
2. test statistic
3. Kolmogorov-Smirnov variable
4. Neyman-Pearson variable

A variable that is a function of the data alone and that can be used to test a hypothesis is called

1. run test
2. test statistic
3. Kolmogorov-Smirnov variable
4. Neyman-Pearson variable

The p-value is the probability

1. that an alternative hypothesis H1 is false
2. of a model being true
3. to observe an equal or larger deviation of the data from a model given the model is true
4. to reject a true hypothesis

The p-value is the probability

1. that an alternative hypothesis H1 is false
2. of a model being true
3. to observe an equal or larger deviation of the data from a model given the model is true
4. to reject a true hypothesis

A difference between the significance level α of a test and the p-value is that

1. the p-value is a random variable while α is not
2. α can be greater than 1 while the p-value cannot
3. α can be always calculated analytically
4. is that the p-value is always greater than 0.9

A difference between the significance level α of a test and the p-value is that

1. the p-value is a random variable while α is not
2. α can be greater than 1 while the p-value cannot
3. α can be always calculated analytically
4. is that the p-value is always greater than 0.9

Suppose a "background only" hypothesis H0 is true and is rejected for a p-value < 0.005. What is the average number of false positive results if 10000 experiments are performed?

1. 0
2. 5
3. 10
4. 50

Suppose a "background only" hypothesis H0 is true and is rejected for a p-value < 0.005. What is the average number of false positive results if 10000 experiments are performed?

1. 0
2. 5
3. 10
4. 50

Testing the goodness of a fit by calculating the maximum deviation of the cumulative distribution function and the corresponding empirical distribution function is known as

1. Neyman-Pearson test

2. Kolmogorov–Smirnov test

3. Wald–Wolfowitz test

4. Gauss-Laplace test

Testing the goodness of a fit by calculating the maximum deviation of the cumulative distribution function and the corresponding empirical distribution function is known as

1. Neyman-Pearson test
2. Kolmogorov–Smirnov test
3. Wald–Wolfowitz test
4. Gauss-Laplace test

Testing a large number of hypotheses about a single data set to find a "significant" effect is sometimes called

1. hypothesis boosting
2. data manipulation
3. p-value hacking
4. type II error

Testing a large number of hypotheses about a single data set to find a "significant" effect is sometimes called

1. hypothesis boosting
2. data manipulation
3. p-value hacking
4. type II error

Let $f(t|H_0)$ be the distribution of a test statistic under hypothesis H_0 and t_{obs} the observed value. The quantity

$$\int_{t_{\mathrm{obs}}}^{\infty} f(t|H_0)\,\mathrm{d}t$$

is called

1. critical region
2. significance level
3. power of the test
4. p-value

Let $f(t|H_0)$ be the distribution of a test statistic under hypothesis H_0 and t_{obs} the observed value. The quantity

$$\int_{t_{\mathrm{obs}}}^{\infty} f(t|H_0)\,\mathrm{d}t$$

is called

1. critical region
2. significance level
3. power of the test
4. p-value

A run test

1. can be used to calculate the chi-squared
2. provides the same information as the chi-squared test
3. tests the hypothesis that the elements of a sequence are mutually independent
4. can be used to calculate the p-value

A run test

1. can be used to calculate the chi-squared
2. provides the same information as the chi-squared test
3. tests the hypothesis that the elements of a sequence are mutually independent
4. can be used to calculate the p-value

The aim of the Bayes factor is to

1. normalize the posterior distribution
2. quantify the support for a model over another
3. construct a credible interval
4. quantify the prior knowledge

The aim of the Bayes factor is to

1. normalize the posterior distribution
2. quantify the support for a model over another
3. construct a credible interval
4. quantify the prior knowledge

Statistical Methods in Particle Physics

Quiz on chapter 7: Confidence Limits and Intervals

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

In Bayesian statistics, an interval in the domain of a posterior probability distribution corresponding to a certain probability is called

1. CLs interval
2. confidence interval
3. confidential interval
4. credible interval

In Bayesian statistics, an interval in the domain of a posterior probability distribution corresponding to a certain probability is called

1. CLs interval
2. confidence interval
3. confidential interval
4. credible interval

In frequentist statistics, the fraction of the time that a confidence interval contains the true value of interest is called

1. coverage
2. credibility
3. power
4. likelihood

In frequentist statistics, the fraction of the time that a confidence interval contains the true value of interest is called

1. coverage
2. credibility
3. power
4. likelihood

The Neyman construction is a

1. Bayesian method
2. Frequentist method
3. Neither a Bayesian nor a frequentist method

The Neyman construction is a

1. Bayesian method
2. Frequentist method
3. Neither a Bayesian nor a frequentist method

You perform one experiment and construct a frequentist confidence interval with a coverage of 68%. The statement that the interval contains the true value with 68% probability is

1. wrong

2. correct

3. ill-defined

You perform one experiment and construct a frequentist confidence interval with a coverage of 68%. The statement that the interval contains the true value with 68% probability is

1. wrong
2. correct
3. ill-defined

"Flip-flopping" refers to a practice of giving upper limits or a confidence interval with lower and upper boundaries depending on the observed result. The problem with this approach is that

1. the upper limit might be negative
2. the coverage of the constructed interval is wrong
3. the confidence interval might be an empty set
4. posterior probability might be alternate between positive and negative values

"Flip-flopping" refers to a practice of giving upper limits or a confidence interval with lower and upper boundaries depending on the observed result. The problem with this approach is that

1. the upper limit might be negative
2. the coverage of the constructed interval is wrong
3. the confidence interval might be an empty set
4. posterior probability might be alternate between positive and negative values

The Feldman-Cousins method is used to

1. determine the confidence level of a given confidence interval
2. avoid the Neyman construction
3. calculate Bayesian credible intervals efficiently
4. construct confidence intervals in the presence of physical boundaries

The Feldman-Cousins method is used to

1. determine the confidence level of a given confidence interval
2. avoid the Neyman construction
3. calculate Bayesian credible intervals efficiently
4. construct confidence intervals in the presence of physical boundaries

The CLs method addresses the problem of

1. spurious exclusions in case of low sensitivity
2. negative upper limits
3. incorrect coverage
4. Bayesian priors

The CLs method addresses the problem of

1. spurious exclusions in case of low sensitivity
2. negative upper limits
3. incorrect coverage
4. Bayesian priors

Statistical Methods in Particle Physics

Quiz on chapter 8: Multivariate analysis

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Please connect to

http://pingo.upb.de/276848

The curve used to quantify the performance of classifiers is called

1. confidence level signature (CLs)
2. performance index measure (PIM)
3. receiver operating characteristic (ROC)
4. Gini index (GI)

The curve used to quantify the performance of classifiers is called

1. confidence level signature (CLs)
2. performance index measure (PIM)
3. receiver operating characteristic (ROC)
4. Gini index (GI)

The naive Bayes classifier is called "naive" because

1. in many situations Bayes' theorem does not apply
2. it approximates PDF's as multi-variate Gaussians
3. it it based on an a linear approximation of Bayes' formula
4. it ignores correlations between the input variables

The naive Bayes classifier is called "naive" because

1. in many situations Bayes' theorem does not apply
2. it approximates PDF's as multi-variate Gaussians
3. it it based on an a linear approximation of Bayes' formula
4. it ignores correlations between the input variables

In supervised learning, "over-training" means that

1. the classifier learns statistical fluctuations of the training sample
2. sometimes classification works better than theoretically expected
3. the classification performance becomes worse when the training sample is too large
4. classification becomes slow for a too large training sample

In supervised learning, "over-training" means that

1. the classifier learns statistical fluctuations of the training sample
2. sometimes classification works better than theoretically expected
3. the classification performance becomes worse when the training sample is too large
4. classification becomes slow for a too large training sample

"Boosting" means

1. to use special relativity in particle identification
2. to combine many weak classifiers into a strong one
3. to increase the performance by using more input variables
4. to use GPUs to speed up the learning process

"Boosting" means

1. to use special relativity in particle identification
2. to combine many weak classifiers into a strong one
3. to increase the performance by using more input variables
4. to use GPUs to speed up the learning process

The Gini index measures

1. the dimension of the feature space
2. the error rate of a classifier
3. the performance of boosted decisions trees
4. the separation between single and background in a sample

The Gini index measures

1. the dimension of the feature space
2. the error rate of a classifier
3. the performance of boosted decisions trees
4. the separation between single and background in a sample