¡article¿
¡presentation¿ Outline
Course Overview
Course Overview
This course is the first part of a two semester introductory course on Mathematical Statistics.
Our goal is to cover Chapters 1-10 of "Mathematical Statistics and Data Analysis", by John A. Rice.
Topics include: Probability, Random Variables, Discrete and Continuous distributions, Order Statistics, Limit Theorems
Roughly speaking, 4450 and 4451 can be broken into two parts:
Math 4450: Probability (mathematics of randomness)
Math 4451: Statistics (procedures for analyzing data)
Course Logistics
About Me (TODO)
Course Website: https://jeswheel.github.io/4450_f25/https://jeswheel.github.io/4450_f25/.
Canvas: Canvas will be used to submit assignments, view grades, and for course announcments.
https://jeswheel.github.io/4450_f25/syllabus.pdfCourse Syllabus
Probability: Chapter I Overview
Probability has been around for a long time.
Probability theory originated in the study of games of chance (i.e., dice, cards, etc.). These provide some nice introduct
More modern examples of probability in practice include:
Modeling mutations in genetics, playing a central role in bioinformatics.
Designing and analyzing computer operating systems.
Modeling atmospheric turbulence.
Probability theory is a cornerstone of the theory of finance, machine learning, and artificial intelligence.
Much more...
This semester will focus on the theory of probability as a mathematical model for chance phenomena. This will be essen
Sample Spaces
Probability theory is concerned with situations in which the outcomes occur randomly. We call these situations experim
Flipping a coin Flipping a coin is an *experiment*, with possible outcomes $\{H, T\}$, which defines the *sample space* of the e
An arbitrary sample space is typically denoted Ω, and an element of Ω is denoted ω.
In the example above, $\Omega = \{H, T\}$.
allowframebreaks]Sample Space Examples Example: Stoplights Driving to work, a commuter passes through a sequence

As statisticians, we want an uncertainty estimate. We want to know how reliable our estimate is, since it is based on on
The data are $y_1^*,\ldots,y_N^*$, which we also write as $y_{1:N}^*$.
Basic estimates of the mean and standard deviation are

This suggests an approximate confidence interval for μ of $\hat\mu_1 \pm 1.96\,\hat\sigma_1/\sqrt{N}$.
Question 1.1. What are the assumptions behind this confidence interval?

fragile]
1955 has missing data, coded as `NA`, requiring a minor modification. So, we compute $\hat\mu_1$ and $\mathrm{SE}_1 = \hat\sigma_1/\sqrt{N}$ as
shadecolorrgb0.969, 0.969, 0.969fgcolor mu1 ¡- mean(y$Low,na.rm=TRUE)se1<-sd(y$Low,na.rm=TRUE)/sqrt(sum(!is

```
mu1 = -2.829268 ,  se1 = 0.6767317
```

Question 1.2. If you had to give an uncertainty estimate on the mean, is it reasonable to present the confidence interval

Some data analysis

The first rule of data analysis is to plot the data in as many ways as you can think of. For time series, we usually start
shadecolorrgb0.969, 0.969, 0.969fgcolor plot(Year,Low,data=y,ty="l")
shadecolorrgb0.969, 0.969, 0.969fgcolor

A first look at an autoregressive-moving average (ARMA) model

ARMA models

Another basic thing to do is to fit an **autoregressive-moving average** (ARMA) model. We'll look at ARMA models

This has a one-lag autoregressive term, $\alpha(Y_{n-1}-\mu)$, and a one-lag moving average term, $\beta\epsilon_{n-1}$. It is therefore called ar
If $\alpha=\beta=0$, we get back to the basic independent model, $Y_n=\mu+\epsilon_n$.
If $\alpha=0$ we have a moving average model with one lag, MA(1).
If $\beta=0$, we have an autoregressive model with one lag, AR(1).
We model $\epsilon_1\ldots,\epsilon_N$ to be an independent, identically distributed (iid) sequence. To be concrete, let's specify a model w

A note on notation

In this course, capital Roman letters, e.g., X, Y, Z, denote random variables. We may also use ϵ, η, ξ, ζ for random noi
We use lower case Roman letters (x, y, z, $\ldots$) to denote numbers. These are not random variables. We use y^* to denote
"We must be careful not to confuse data with the abstractions we use to analyze them." (William James, 1842-1910).
Other Greek letters will usually be parameters, i.e., real numbers that form part of the model.

Fitting an ARMA model in R

fragile]Maximum likelihood

We can readily fit the ARMA(1,1) model by maximum likelihood,
shadecolorrgb0.969, 0.969, 0.969fgcolor arma11 ¡- arima(y$Low,order=c(1,0,1))
`print(arma11)` or just `arma11` gives a summary of the fitted model, where α is called `ar1`, β is called `ma1`, and μ is call
shadecolorrgb0.969, 0.969, 0.969fgcolor

```
Coefficients:
         ar1    ma1  intercept
      -0.584  0.619     -2.823
s.e.   0.598  0.578      0.688

sigma^2 estimated as 55.8:  log likelihood = -421.84,
aic = 851.68
```

We will write the ARMA(1,1) estimate of μ as $\hat\mu_2$, and its standard error as SE_2.

fragile]Investigating R objects

Some poking around is required to extract the quantities of primary interest from the fitted ARMA model in R.
shadecolorrgb0.969, 0.969, 0.969fgcolor names(arma11)

```
 [1] "coef"      "sigma2"    "var.coef"  "mask"      "loglik"
 [6] "aic"       "arma"      "residuals" "call"      "series"
[11] "code"      "n.cond"    "nobs"      "model"
```

shadecolorrgb0.969, 0.969, 0.969fgcolor mu2 ¡- arma11$coef["intercept"]se2<-sqrt(arma11$var.coef["intercept","interce

```
mu2 = -2.823284 ,  se2 = 0.6880817
```

Model diagnostics

fragile]Comparing the iid estimate with the ARMA estimate

In this case, the two estimates, $\hat\mu_1=-2.83$ and $\hat\mu_2=-2.82$, and their standard errors, $\mathrm{SE}_1=0.68$ and $\mathrm{SE}_2=0.69$, are
For data up to 2015, $\hat\mu_1^{2015}=-2.83$ and $\hat\mu_2^{2015}=-2.85$, with standard errors, $\mathrm{SE}_1^{2015}=0.68$ and $\mathrm{SE}_2^{2015}=0.83$.
In this case, the standard error for the simpler model is $100(1-\mathrm{SE}_1^{2015}/\mathrm{SE}_2^{2015}) = 17.5\%$ smaller.
Exactly how the ARMA(1,1) model is fitted and the standard errors computed will be covered later.
Question 1.3. When standard errors for two methods differ, which is more trustworthy? Or are they both equally valid