Statistics and Data Analysis
Unit 06 – Lecture 06 Notes
ADF Test for Stationarity

Tofik Ali

February 17, 2026

Topic

ADF test (unit root) and interpretation.

How to Use These Notes

These notes are written for students who are seeing the topic for the first time. They follow the slide order, but add the missing 'why', interpretation, and common mistakes. If you get stuck, look at the worked exercises and then run the Python demo.

Course repository (slides, demos, datasets): `https://github.com/tali7c/Statistics-and-Data-Analysis`

Time Plan (55 minutes)

- 0–10 min: Attendance + recap of previous lecture
- 10–35 min: Core concepts (this lecture's sections)
- 35–45 min: Exercises (solve 1–2 in class, rest as practice)
- 45–50 min: Mini demo + interpretation of output
- 50–55 min: Buffer / wrap-up (leave 5 minutes early)

Slide-by-slide Notes

Title Slide

State the lecture title clearly and connect it to what students already know. Tell students what they will be able to do by the end (not just what you will cover).

Quick Links / Agenda

Explain the structure of the lecture and where the exercises and demo appear.

- Overview

- ADF Test
- Interpretation
- Exercises
- Demo
- Summary

Learning Outcomes

- State null and alternative of ADF test (unit root)
- Interpret ADF p-value for stationarity decision
- Apply ADF to original and differenced series (idea)
- Explain why tests are not the only evidence (plots matter)

Why these outcomes matter. **Degrees of freedom (df)** roughly represent how much independent information is available to estimate variability. For a one-sample t-test, df $= n - 1$ because one constraint is used to estimate the sample mean. df affects the critical values and the shape of the t-distribution (small df -> heavier tails). A **p-value** is computed assuming the null hypothesis H_0 is true. It measures how surprising the observed data (or something more extreme) would be under H_0. A small p-value suggests the data is hard to explain by H_0 alone, but it does not tell you how large the effect is or whether it is practically important.

ADF Test: Key Points

- H0: unit root (non-stationary)
- H1: stationary
- Small p-value -> reject H0

Explanation. The **null hypothesis** H_0 usually represents 'no effect' or a baseline value (e.g., $\mu = 60$). The **alternative** H_1 represents the effect you are looking for (e.g., $\mu \neq 60$ or $\mu > 60$). We compute a test statistic and a p-value assuming H_0 is true. **Degrees of freedom (df)** roughly represent how much independent information is available to estimate variability. For a one-sample t-test, df $= n - 1$ because one constraint is used to estimate the sample mean. df affects the critical values and the shape of the t-distribution (small df -> heavier tails). A **p-value** is computed assuming the null hypothesis H_0 is true. It measures how surprising the observed data (or something more extreme) would be under H_0. A small p-value suggests the data is hard to explain by H_0 alone, but it does not tell you how large the effect is or whether it is practically important.

Interpretation: Key Points

- If non-stationary, difference and test again
- Seasonality can require seasonal differencing
- Use ACF/PACF + diagnostics too

Explanation. **Seasonality** is a repeating pattern with a fixed period (weekly, monthly, yearly). You must account for it; otherwise forecasts systematically miss repeating rises/falls. Seasonal differencing and SARIMA are common tools. **Differencing** transforms a series by subtracting the previous value: $y_t - y_{t-1}$. It removes trend and can help achieve stationarity. Over-differencing can add noise, so use the smallest differencing order that works. **Stationarity** (intuition) means the series behavior is stable over time: roughly constant mean/variance and correlation structure. AR/MA/ARIMA models assume stationarity (after differencing). If the process changes over time, parameters learned from the past may not hold.

Exercises (with Solutions)

Attempt the exercise first, then compare with the solution. Focus on interpretation, not only arithmetic.

Exercise 1: ADF null

What is H0 in ADF?

Solution

- Unit root; non-stationary.

Walkthrough. The **null hypothesis** H_0 usually represents 'no effect' or a baseline value (e.g., $\mu = 60$). The **alternative** H_1 represents the effect you are looking for (e.g., $\mu \neq 60$ or $\mu > 60$). We compute a test statistic and a p-value assuming H_0 is true. **Degrees of freedom (df)** roughly represent how much independent information is available to estimate variability. For a one-sample t-test, df $= n-1$ because one constraint is used to estimate the sample mean. df affects the critical values and the shape of the t-distribution (small df -> heavier tails).

Exercise 2: Decision

If p=0.02 at alpha=0.05, what do you conclude?

Solution

- Reject H0; evidence of stationarity.

Walkthrough. The **null hypothesis** H_0 usually represents 'no effect' or a baseline value (e.g., $\mu = 60$). The **alternative** H_1 represents the effect you are looking for (e.g., $\mu \neq 60$ or $\mu > 60$). We compute a test statistic and a p-value assuming H_0 is true. The **significance level** α is the maximum Type I error rate you are willing to tolerate: the probability of rejecting H_0 when H_0 is actually true. Common choices are 0.05 or 0.01, but the right value depends on consequences of false alarms vs missed detections.

Exercise 3: Next step

If p=0.6, what next step?

Solution

- Difference and test again; consider seasonal differencing.

Walkthrough. **Seasonality** is a repeating pattern with a fixed period (weekly, monthly, yearly). You must account for it; otherwise forecasts systematically miss repeating rises/falls. Seasonal differencing and SARIMA are common tools. **Differencing** transforms a series by subtracting the previous value: $y_t - y_{t-1}$. It removes trend and can help achieve stationarity. Over-differencing can add noise, so use the smallest differencing order that works.

Mini Demo (Python)

Run from the lecture folder:

```
python demo/demo.py
```

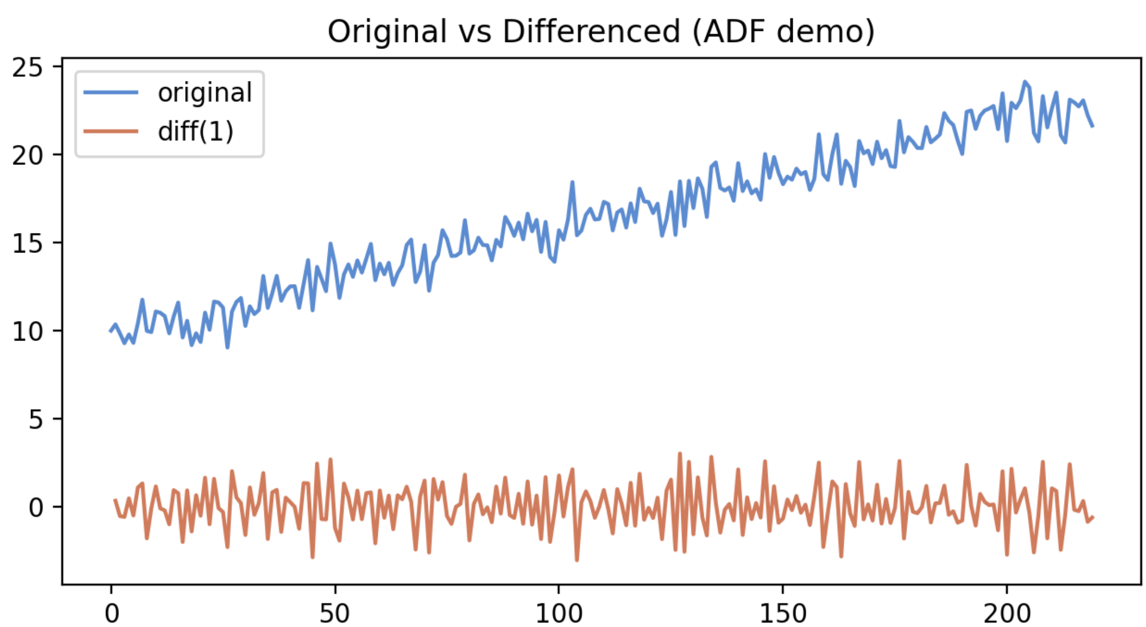
Output files:

- `images/demo.png`
- `data/results.txt`

What to show and say.

- Runs an ADF test on an example series and reports the p-value.
- Shows how conclusions can change after differencing.
- Use it to stress: combine tests with plots/ACF/PACF, not tests alone.

Demo Output (Example)



Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

Exit Question

Why should we not rely on only one test to decide stationarity?

Suggested answer (for revision). Plots and ACF/PACF provide context; a single test can be noisy/misleading under seasonality, breaks, or short samples.

References

- Montgomery, D. C., & Runger, G. C. *Applied Statistics and Probability for Engineers*, Wiley.
- Devore, J. L. *Probability and Statistics for Engineering and the Sciences*, Cengage.
- McKinney, W. *Python for Data Analysis*, O'Reilly.

Appendix: Slide Deck Content (Reference)

The material below is a reference copy of the slide deck content. Exercise solutions are explained in the main notes where applicable.

Title Slide

https://github.com/tali7c/Statistics-and-Data-Analysis

Quick Links

Overview **ADF Test** **Interpretation** **Exercises** **Demo** **Summary**

Agenda

- Overview
- ADF Test
- Interpretation
- Exercises
- Demo
- Summary

Learning Outcomes

- State null and alternative of ADF test (unit root)
- Interpret ADF p-value for stationarity decision
- Apply ADF to original and differenced series (idea)
- Explain why tests are not the only evidence (plots matter)

ADF Test: Key Points

- H0: unit root (non-stationary)
- H1: stationary
- Small p-value -> reject H0

Interpretation: Key Points

- If non-stationary, difference and test again
- Seasonality can require seasonal differencing
- Use ACF/PACF + diagnostics too

Exercise 1: ADF null

What is H0 in ADF?

Solution 1

- Unit root; non-stationary.

Exercise 2: Decision

If p=0.02 at alpha=0.05, what do you conclude?

Solution 2

- Reject H0; evidence of stationarity.

Exercise 3: Next step

If p=0.6, what next step?

Solution 3

- Difference and test again; consider seasonal differencing.

Mini Demo (Python)

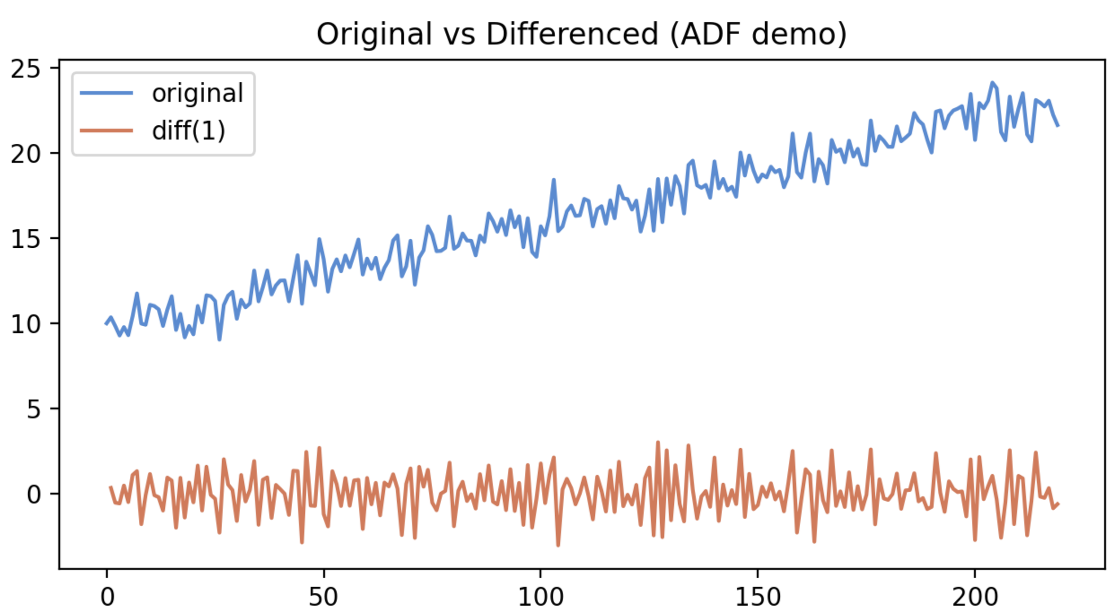
Run from the lecture folder:

```
python demo/demo.py
```

Outputs:

- `images/demo.png`
- `data/results.txt`

Demo Output (Example)



Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

Exit Question

Why should we not rely on only one test to decide stationarity?