Universitas Gadjah Mada
Fakultas Teknik
Departemen Teknik Sipil dan Lingkungan
Prodi Magister Teknik Sipil

Statistika Teknik

Pendahuluan

Statistika

- Statistika/metode statistis
  - adalah metode pengolahan data yang didapat dari suatu operasi berulang-ulang.
- Operasi dilakukan melalui
  - observasi
  - pengukuran
  - experimen

Pemakaian Statistika

- *Descriptive statistics*
- *Statistical inference*

Descriptive Statistics, Statistical Inference

- Contoh
  - Hasil wawancara terhadap responden yang dipilih secara acak dari sebuah wilayah, menunjukkan bahwa 15% responden merupakan pelanggan PDAM.
  - Kemudian dilakukan kampanye pemakaian air PDAM kepada penduduk wilayah tersebut.
  - Hasil wawancara terhadap responden yang dilakukan beberapa waktu kemudian setelah kampanye tersebut menunjukkan bahwa 21% responden merupakan pelanggan PDAM.

Descriptive Statistics, Statistical Inference

- Pertanyaan
  - Apakah kita benar dalam menganggap bahwa kenaikan jumlah pelanggan dari 15% pada kelompok responden pertama menjadi 21% pada kelompok responden kedua adalah akibat kampanye pemakaian air PDAM?
  - Apabila kita meragukan bahwa kampanye tersebut telah meningkatkan jumlah pelanggan, bagaimanakah dengan data hasil wawancara tersebut?
  - Bagaimana dengan data dari penduduk yang tidak diwawancarai?

Descriptive Statistics, Statistical Inference

- Descriptive statistics
  - Angka 15% dipakai untuk menunjukkan bagian dari penduduk pada responden pertama yang memakai air PDAM.
  - Dalam hal ini, angka 15% tersebut merangkum dengan ringkas, padat, dan jelas fakta yang tak diperlihatkan bahwa dari 140 responden, 21 adalah pemakai air PDAM.
  - Dengan demikian, dapat dikatakan bahwa
    - descriptive statistics adalah pemakaian angka untuk merangkum informasi yang diketahui mengenai suatu keadaan.

Descriptive Statistics, Statistical Inference

- Statistical inference
  - Apabila kita memakai hasil wawancara tadi untuk mengatakan bahwa sekitar 15% penduduk wilayah tersebut adalah pemakai air PDAM, maka
  - kita memakai angka untuk menyimpulkan suatu hal mengenai populasi, yang lebih besar daripada responden (sampel), yang kita sendiri tidak memiliki informasi yang lengkap tentang populasi tersebut.
  - Dengan demikian dapat dikatakan bahwa
    - statistical inference adalah pemakaian angka untuk mengatakan suatu informasi mengenai populasi, yang pada umumnya lebih besar daripada sampel darimana data diperoleh.

Contoh Pemakaian Statistika

- Contoh #1
  - Sebutkan beberapa contoh pemakaian statistik yang Saudara kenal.
  - Diskusikan
    - *descriptive statistics*,
    - *statistical inference*.

Contoh Pemakaian Statistika

- Contoh #2
  - Misal dijumpai suatu penyakit dimana ¾ penderitanya sembuh dalam 3 bulan tanpa pengobatan.
  - Kemudian seorang dokter menyatakan bahwa dia telah menemukan obat penyembuh penyakit tersebut.
  - Obat tersebut diberikan kepada 100 penderita.
  - Kalaupun obat tersebut tak berfungsi, kita tetap dapat memperkirakan bahwa 75 penderita akan sembuh.

Contoh Pemakaian Statistika

- Contoh penyakit vs obat
  - Salah satu permasalahan dalam inferensi statistik dalam contoh di atas adalah:
    - berapa jumlah penderita yang harus sembuh sebelum kita menerima bahwa obat penemuan dokter tersebut benar-benar menyembuhkan penyakit tersebut.
    - Tentu saja, apabila ke-100 penderita tersebut sembuh, maka kita akan dengan sangat antusias menerima bahwa obat tersebut benar-benar menyembuhkan.
    - Namun, apabila jumlah penderita yang sembuh adalah 95, atau 90, atau 85, apakah kita bisa menerima klaim tersebut?

Contoh Pemakaian Statistika

- Jadi?
  - Tugas memutuskan atau menjawab pertanyaan tersebut merupakan satu hal penting bagi seorang statistikawan.
- Diskusi
  - Dapatkah kita mengatakan dengan yakin bahwa obat tersebut benar-benar manjur?
  - Bahkan andaikata ke-100 penderita tersebut sembuh, masih ada kemungkinan (walaupun kecil) bahwa ke-100 penderita tersebut memang sembuh dengan sendirinya, tanpa ada kontribusi dari obat.
    - Ada kemungkinan bahwa ke-100 penderita memang memiliki kekebalan terhadap penyakit tersebut.

Contoh Pemakaian Statistika

- Peringatan
  - Adalah penting untuk diperhatikan dalam pengambilan sampel (penderita, dalam hal ini):
    - bahwa sampel dipilih tanpa bias,
    - bahwa sampel benar-benar mencerminkan situasi populasi yang kita ingin ketahui perilakunya (dalam hal ini: para penderita penyakit tersebut).

Statistika di Bidang Keairan

- Modeling
  - Permasalahan keairan (hidraulika, dan khususnya hidrologi) umumnya dikaji dan dicari solusinya melalui pemakaian suatu model.

Statistika di Bidang Keairan

- Jenis model
  - Persamaan empirik
  - Model fisik
  - Model matematik (persamaan matematik, statistik, dan empirik)

Statistika di Bidang Keairan

- Model hidraulika dan (khususnya) hidrologi dapat dibedakan menjadi
  - Model deterministik
  - Model parametrik
  - Model stokastik

Model Hidrologi

- Model deterministik
  - Pasti
  - Memakai persamaan yang didasarkan pada proses fisik
  - Tidak membutuhkan data experimen untuk menerapkannya
  - Model diterapkan berulang-ulang dengan input yang sama selalu memberikan output yang sama
- Model stokastik
  - Tidak pasti, perkiraan, probababilistik
  - Parameter model ditetapkan (diperkirakan) berdasarkan data pengamatan atau pengukuran
  - Model diterapkan berulang-ulang dengan input yang sama dapat memberikan output yang berbeda

Model Hidrologi

- Model parametrik
  - Parameter model ditetapkan sekali, kemudian dipakai terus selama data pengamatan (yang dipakai untuk menetapkan parameter model tsb) tidak berubah

Model & Data

- Penetapan/perhitungan parameter model harus didasarkan pada data pengamatan/pengukuran
  - Kualitas model (valid dan dapat diterapkan) bergantung pada kualitas data yang dipakai untuk menetapkan parameter model.

Sebaik-baik kualitas model, tidak akan melebihi kualitas data yang dipakai.

Data

- Contoh data
  - Hidrologi (misal pengukuran di stasiun meteorologi)
    - curah hujan
    - penguapan
    - resapan

Data

- Contoh data
  - Hidraulika (misal experimen di lab)
    - kecepatan aliran
    - tekanan
    - volume sedimen

Data: Homogen dan Representatif

- Data untuk menetapkan parameter model harus memenuhi syarat:
  - homogen, dan
  - representatif

Data

- Contoh data
  - Danau dan pantai (misal pengukuran di laut)
    - kecepatan aliran
    - kecepatan angin
    - tinggi, arah, periode gelombang

Data: Homogen dan Representatif

- Penyebab data tidak homogen
  - stasiun pengukuran berpindah
  - aliran (sungai) berpindah
  - pembangunan struktur baru (*dam*, *reservoir*)
  - perbaikan alur sungai
  - perubahan land-use

Data: Homogen dan Representatif

- Penyebab data tidak representatif
  - data diambil hanya pada musim tertentu
  - jumlah data tidak cukup banyak (seri data pendek)

Sampel dan Populasi

- Apakah sampel itu?
  - Sampel adalah kumpulan objek (data) yang diambil dari kumpulan yang lebih besar dari objek tersebut.
  - Kumpulan yang lebih besar tersebut, apabila mengandung semua objek yang mungkin (*all objects possible*), disebut populasi.

Sampel dan Populasi

- Apakah sampel itu?
  - Berdasarkan sampel, kita dapat menarik kesimpulan mengenai populasi darimana sampel tersebut diperoleh.
  - Statistik adalah salah satu alat yang memungkinkan kita dapat menarik kesimpulan yang valid mengenai populasi berdasarkan informasi yang terkandung di dalam sampel.

Random Sample

- Sampel random (acak)
  - Apabila setiap elemen populasi memiliki probabilitas (peluang) yang sama untuk terambil kedalam sampel, atau
  - Sampel yang dipilih dengan cara sedemikian hingga sampel lain yang diambil dari populasi yang sama akan memiliki karakteristik yang sama.

Kesalahan Data

- Kesalahan data
  - Kesalahan pengukuran
  - Kesalahan transmisi data
  - Kesalahan pengolahan data
- Sifat kesalahan data
  - Kesalahan sistematik
  - Kesalahan random

Kesalahan Sampel

- Kesalahan sampel (*sampling errors*)
  - Tidak sama arti dengan kesalahan data, namun lebih menunjukkan pada variabilitas sampel atau ketidak-pastian sampel.
  - Kesalahan sampel merupakan sifat yang melekat pada sampel random.

Kesalahan Sampel

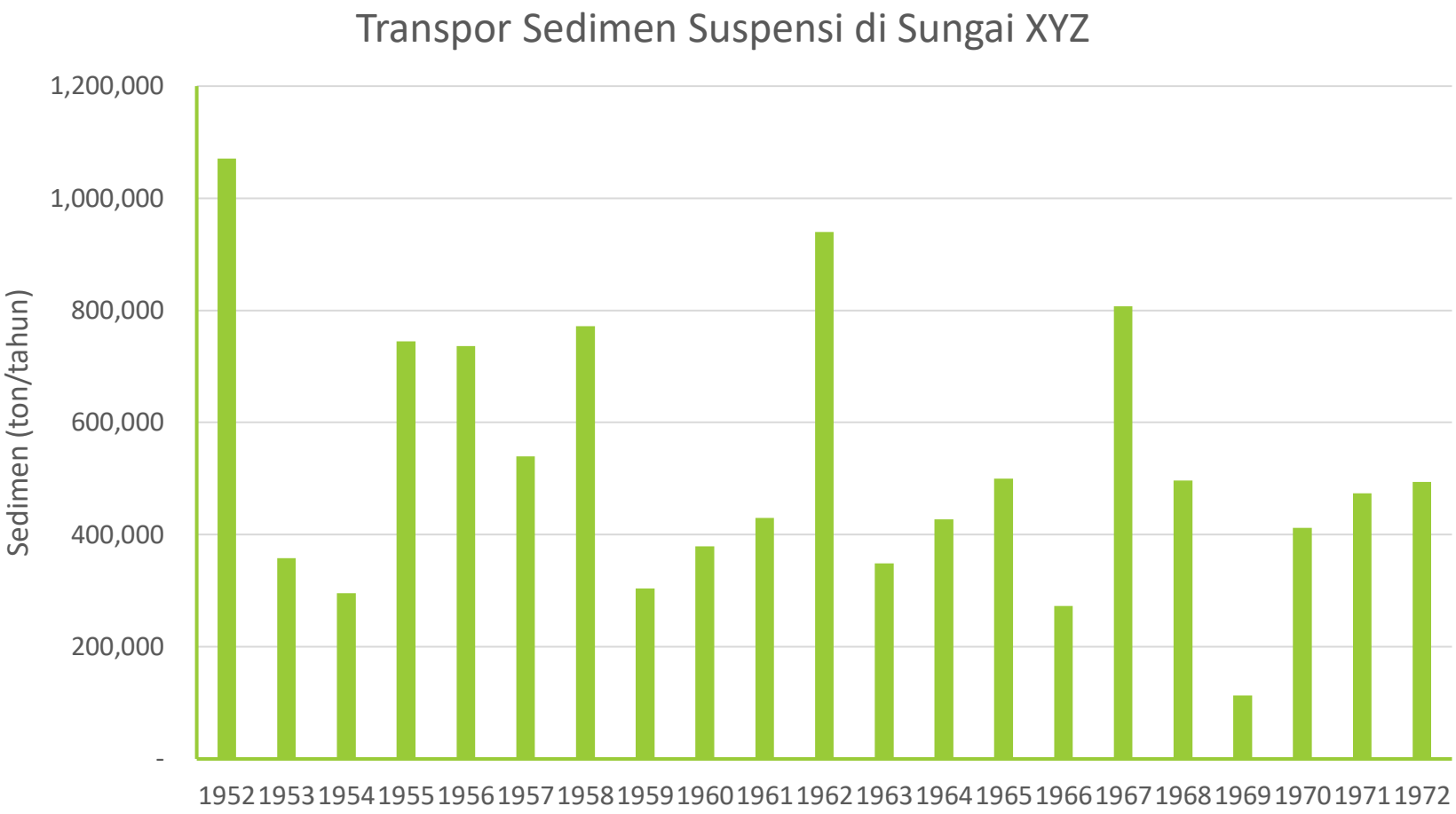
- Kesalahan sampel (*sampling errors*)
  - Suatu sampel random memiliki karakteristik/sifat statistik yang mirip dengan yang dimiliki populasinya dan akan sama dengan yang dimiliki populasinya hanya apabila ukuran sampel sangat besar (atau sampel mencakup seluruh populasi).
  - Contoh, dua sampel yang diambil dari satu populasi yang sama, maka karakterisktik statistik kedua sampel akan mirip, namun akan sama hanya apabila ukuran kedua sampel sangat besar.

Transpor Sedimen Suspensi di Sungai XYZ

Tahun	Sed. susp. (ton)	Sed. susp. rata-rata (ton)	
1952	1,070,886		
1953	357,517		
1954	295,187	640,827	
1955	744,412		
1956	736,131		
1957	539,645		
1958	771,530		
1959	303,815	484,739	
1960	378,891		
1961	429,815		
1962	940,227		519,611
1963	348,653		
1964	426,989	497,604	
1965	499,532		
1966	272,618		
1967	807,457		
1968	496,158		
1969	112,737	460,392	
1970	411,824		
1971	473,785		
1972	494,030		

Kesalahan Data *vs* Kesalahan Sampel

- Contoh ilustrasi
  - data transpor sedimen suspensi suatu sungai
  - satuan ton → ton/tahun
  - tahun 1952 – 1972


Transpor Sedimen Suspensi di Sungai XYZ
Sedimen (ton/tahun)
1,200,000
1,000,000
800,000
600,000
400,000
200,000
-
1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972

Kesalahan Data *vs* Kesalahan Sampel

- *Systematic errors*
  - sampling dilakukan hanya saat kedalaman aliran melampaui nilai tertentu yang sudah ditetapkan sebelumnya,
  - shg. tidak dilakukan pada saat *low flow,*
  - padahal konsentrasi sedimen pada saat low flow biasanya lebih kecil daripada saat debit besar,
  - dengan demikian, terjadi *built in error*.

Kesalahan Data *vs* Kesalahan Sampel

- *Measurement errors*
  - peralatan tersumbat,
  - penempatan peralatan (misal Delft Bottle) yang tidak sejajar aliran,
  - transpor sedimen dasar (*bed load*) ikut terambil,
  - dll.

Kesalahan Data *vs* Kesalahan Sampel

- *Data transmittal and processing errors*
  - kesalahan pada saat menuliskan data dari formulir survey ke file data,
  - salah tulis, salah ketik,
  - dll.

Kesalahan Data *vs* Kesalahan Sampel

- Sampling errors
  - Anggap bahwa tidak terjadi kesalahan data (tak ada systematic, measurement, transmittal or processing errors).
  - Sedimen suspensi rata-rata tahunan dalam lima tahun berurutan:
    - 640.827 ton
    - 484.739 ton
    - 497.604 ton
    - 460.392 ton

BEDA

Kesalahan Data *vs* Kesalahan Sampel

- *Sampling errors*
  - Walau tanpa kesalahan data, diperoleh 4 nilai berbeda.
  - Setiap nilai tidak mengandung kesalahan data, namun keempatnya menunjukkan nilai berbeda.
  - Kesalahan ini (*sampling errors*) diakibatkan oleh adanya variabilitas fenomena (transpor sedimen) yang di-sampling.

Kesalahan Data *vs* Kesalahan Sampel

- Meminimumkan *sampling errors*
  - Menambah jumlah sampel

Kala Ulang

- Kala ulang (*return period*) hidrologis
  - Antar-waktu rata-rata antar terjadinya peristiwa (hidrologis) dengan besaran tertentu yang disamai atau dilampaui.
    - Contoh: debit 25-tahunan adalah debit yang disamai atau dilampaui rata-rata sekali dalam 25 tahun **selama kurun waktu yang cukup (sangat) panjang**.

Kala Ulang

- Kala ulang (*T*) dan probabilitas suatu peristiwa terlampaui (*p*)
  - Jika suatu peristiwa, yang memiliki besaran tertentu, terlampaui terjadi rata-rata sekali dalam kurun waktu 25 tahun, maka probabilitas peristiwa tersebut terjadi dalam satu tahun adalah 1/25 = 0,04 atau 4%.

$$T = \frac{1}{p}$$

Terima kasih