CALVIN
INSTITUTE OF TECHNOLOGY
God's People for God's Glory

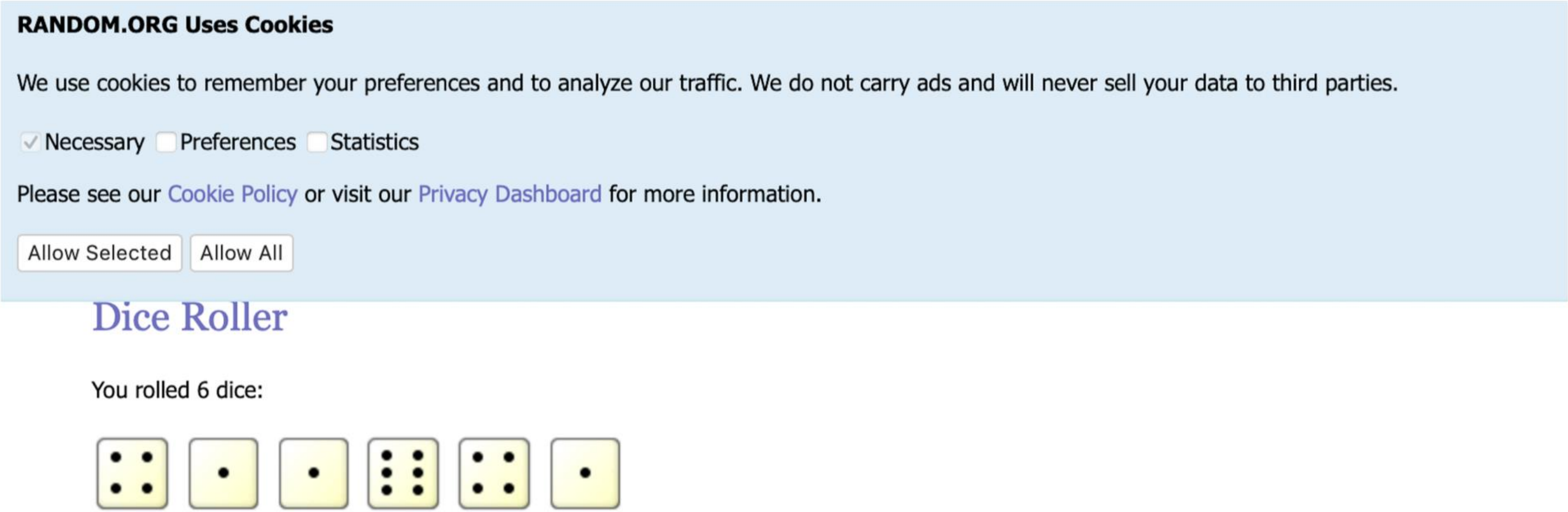
Ekspektasi: Teori dan Fakta

Konsep Dasar Hukum Bilangan Besar

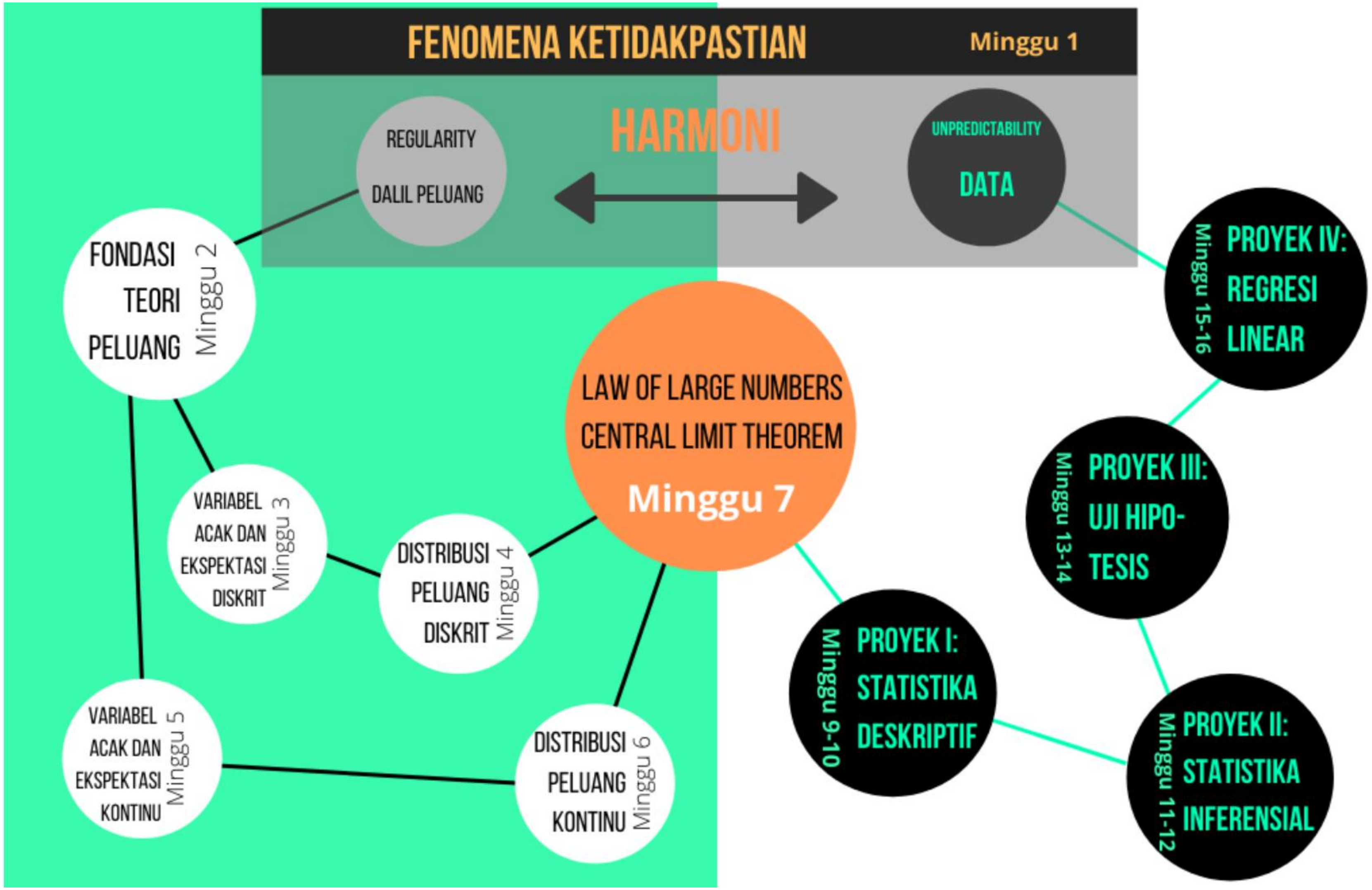
1.1 Ekspektasi: Teori dan Kenyataannya

Misalkan X menyatakan banyaknya mata dadu 1 yang muncul dari pelemparan enam dadu. Secara teori, kita tahu bahwa $X \sim \text{Bin}(6, \frac{1}{6})$ dengan ekspektasi $E[X] = 1$. Bagaimana dengan kenyataanya?

- Lakukan simulasi pada link: `https://www.random.org/dice/?num=6`.
- Lengkapi `https://bit.ly/SimulasiDaduRekap` dengan hasil simulasi dadu.



Kita akan melihat kaitan antara teori peluang dan statistika sebagai gambaran besar mata kuliah ini.



CALVIN INSTITUTE OF TECHNOLOGY

1.2 Hukum Bilangan Besar

Teorema. Misalkan X variabel acak dengan nilai ekspektasi atau rata-rata ____. Misalkan pula diberi data empiris $X_1, X_2, \ldots, X_n$ adalah hasil ____________________ dari variabel acak X yang dilakukan secara __________________________________ (*identically & independently distributed*):

- identik artinya $X_1 \sim X, X_2 \sim X, \ldots, X_n \sim X$,
- saling bebas artinya X_i tidak mempengaruhi X_j untuk setiap $i \neq j$.

Maka, nilai rata-rata empiris $\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n) \to \mu_X$ saat $n \to \infty$.

Misalkan X variabel acak yang menyatakan banyak sisi 1 dari pelemparan enam buah dadu.

Teori	Kenyataan
$\mu_X =$ ______	• $\bar{X}_{10} =$ ______ • $\bar{X}_{100} =$ ______ • $\bar{X}_{1000} =$ ______

1.3 Populasi dan Sampel

Secara rata-rata, dalam satu minggu, berapa kali orang indonesia pergi keluar rumah?

- Misalkan X variabel acak banyaknya orang yang pergi keluar rumah dalam 1 minggu.
- μ_X sebagai rata-rata dari X, yakni $\frac{1}{N}(X_1 + X_2 + \ldots + X_N)$.
- Populasi yang menjadi fokus perhatian adalah WNI, namun ini tidak mudah jika dilakukan untuk $N = |\text{Populasi}| \approx 250$ juta orang. Oleh karena itu, kita perlu mengambil $k = |\text{Sampel}| \ll N$.
- Nilai statistik dari ______________ biasanya lebih sulit untuk dihitung nilainya.
- Nilai statistik dari ______________ biasanya lebih mudah untuk dihitung nilainya. Nilai statistik ini dapat digunakan untuk memberikan gambaran atau estimasi nilai statistik ____________________.

Definisi. Misalkan suatu populasi memiliki rata-rata ______. Misalkan pula diambil ______ buah sampel dari populasi yang diambil secara _____________________ dan ___________________________ dengan hasil realisasi $x_1, x_2, \ldots, x_N$. Maka, rata-rata dari sampel, yakni

$$\bar{x}_N = \frac{1}{N}(x_1 + x_2 + \ldots + x_N),$$

disebut sebagai ____ bagi μ_X.

Kalimat "$\bar{x}_N$ adalah estimator titik bagi μ_X" memiliki beberapa intepretasi.

- Nilai persis dari ______ tidak diketahui, sedangkan nilai dari ______ dapat diketahui atau dihitung.
- Berdasarkan _____________________, nilai _____________________ ketika _____________________.
- Semakin banyak sampel yang diambil, maka semakin akurat ______ untuk mengestiamasi ______.

CALVIN
INSTITUTE OF TECHNOLOGY
God's People for God's Glory

Konsep Populasi dan Sampel

Teorema Limit Pusat dan Estimasi Rataan

1.4 Teorema Limit Pusat

Sebagai ilustrasi, beberapa masalah dalam kehidupan sehari-hari yang memiliki distribusi beragam:

- data tinggi badan mahasiswa CIT,
- data harga saham harian,
- data persentase warga setuju suatu kebijakan pemerintah,
- data banyakanya mahasiswa per angkatan,
- data hasil pelemparan dadu,
- data pendapatan harian pengemudi ojek online.

Bagaimana kita mengetahui seberapa $\bar{x}_N$ dekat atau jauh dari μ_X?

Teorema. Misalkan suatu populasi atau variabel acak ________ memiliki rata-rata ________ dan variansi ________. Misalkan pula ________ data sampel diambil dari populasi secara ____________________, yakni $x_1, x_2, \ldots, x_N \sim X$. Jika ______ adalah rata-rata dari sampel, maka

$$Z = \frac{\bar{x}_N - \mu_X}{\sigma_X/\sqrt{N}} \sim \mathcal{N}(0,1),$$

ketika $N \to \infty$. Biasanya, kita gunakan $N \geq 30$ sebagai kondisi 'tak hingga' ini. Lebih jauh, konvergensi ini berlaku tanpa mempedulikan ______________________ dari variabel acak X.

1. **Estimasi rataan sampel dari populasi.** Waktu tempuh MRT dari stasiun Lebak Bulus ke stasiun Bundaran HI diketahui secara umum memiliki rata-rata 32 menit dengan standar deviasi 5 menit. Dalam suatu hari, diketahui MRT beroperasi bolak-balik dalam jalur ini sebanyak 50 kali. Dalam satu hari ini, berapa peluang rata-rata lama perjalanannya lebih dari 35 menit?

2. **Estimasi rataan populasi dari sampel.** Waktu proses pembuatan KTP bervariasi secara acak tanpa diketahui nilai rata-ratanya. Namun, diketahui bahwa variansinya adalah 1 hari. Anda mengumpulkan data dari 50 teman Anda mengenai lama pembuatan KTP dari pengalaman mereka. Setelah dirata-rata, diperoleh nilai $4,5$ hari.

   (a) Seberapa mungkin bahwa rata-rata lama proses pembuatan KTP berkisar antara $4-5$ hari?

(b) Tentukan nilai k terkecil sehingga Anda dapat yakin dengan tingkat kepercayaan 95% bahwa secara umum, rata-rata pembuatan KTP tidak melebihi k hari. Jelaskan.