1) Exponential Family & Sufficiency

🧀 01 Poisson Meets the Exponential Family

Let X be a random variable following a Poisson distribution with parameter $\lambda > 0$, i.e.,
$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ for $k = 0, 1, 2, \dots$

- a. Show that the Poisson distribution belongs to the exponential family by writing its PMF in the form

$$f(x \mid \lambda) = h(x)\exp(\eta(\lambda)T(x) - A(\lambda)).$$

  Identify $h(x), \eta(\lambda), T(x)$, and $A(\lambda)$.
- b. Using the exponential family form, what is the sufficient statistic for λ based on an i.i.d. sample $X_1, \dots, X_n$?

$\eta(\lambda)\ T_x - A\lambda$

$$P(X=k) = \frac{\lambda^{x} e^{-\lambda}}{x!} = \frac{e^{x\ln\lambda}\cdot e^{-\lambda}}{x!} = \frac{1}{x!}\exp(x\ln\lambda - \lambda)$$

$$h(x) = \frac{1}{x!}$$

$$f(x \mid \lambda) = \prod_{i=1}^{n} \frac{1}{x_i!}\exp\left(\ln\lambda \sum_{i=1}^{n} x_i - n\lambda\right)$$

$T(x)$

$p = 0.7$ $\quad \boxed{\{1, 0, 0, 1, \dots\}}$ $\quad 1$

$$f(x \mid \theta) = g(T(x), \theta)\, h(x)$$

$$\prod e^{\{\dots\}}$$

$$\boxed{T(x) = \sum_{i=1}^{n} X_i}$$

$$\bar{X} = \frac{T}{n} \qquad g(x) = \frac{x}{n}$$

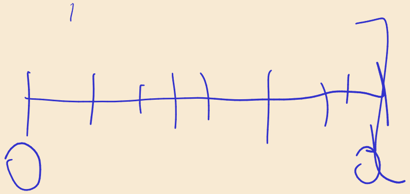
$$g(T(x))$$

🧀 🧀 02 Slit Width Estimation

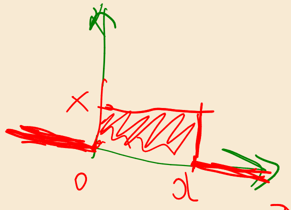
In an experiment, n drops of solution are released uniformly through a slit onto a surface. We model the one-dimensional impact points $X_1, \dots, X_n$ as i.i.d. $\mathrm{Uniform}(0, d)$, where the unknown slit width $d > 0$ is to be estimated.

- a. Write down the joint density $f(\mathbf{x} \mid d)$ for the sample.
- b. Using the Fisher–Neyman factorization theorem, show that $X_{(n)} = \max\{X_1, \dots, X_n\}$ is sufficient for d.
- c. Is $X_{(n)}$ unbiased for d? If not, find an unbiased estimator based on $X_{(n)}$.

Hint for (c): the CDF of $X_{(n)}$ is $F_{X_{(n)}}(x) = (x/d)^n$ for $0 \le x \le d$.



a, b $\quad \boxed{\frac{b+a}{2}}$



$d>0$

$X_i \sim \mathcal{U}(0, d)$

$d-0 \cdot x = 1$

$x = \frac{1}{d-0}$

$\frac{1}{b-a}$

$1, 2.5, 3, 2$

$f(x \mid d) = \prod^{n} \frac{1}{d} \boxed{\mathbb{1}(0 < x_i < d)} =$

$\hat{\theta} = \frac{1}{d^n} \mathbb{1}(X_{(1)} \geq 0) \, \mathbb{1}(X_{(n)} \leq d)$

$g(X_{(n)}, d) \quad h(x)$

$\boxed{E[\hat{\theta}] = d}$

$\hat{\theta}_1 \to 1000$

$\hat{\theta}_2 \to 2000$

$1, 2, 3 \mapsto \boxed{3}$

$1, 4, 4, 5 \to 4.5$

$2, 1, 4, 3 \to 4$

$\frac{3+4.5+4}{3} = 5$

$$E[X(n)] = \int_0^d x \cdot$$

$$\frac{n-1}{n}\left(\frac{x^n}{d}\right)$$

$$\int_0^d x \frac{n-1}{n} \, n \frac{x^{n-1}}{d^n} = \boxed{\frac{n}{n+1} d}$$

bias

$$T(x) = X_{(n)} \cdot \frac{n+1}{n}$$

$0 \quad d$

$1, 2, \boxed{3}, 2.5, 2$

$3 \cdot \frac{6}{5} \to d$

🧀 🧀 04 Binomial Sufficiency and Estimating π^2

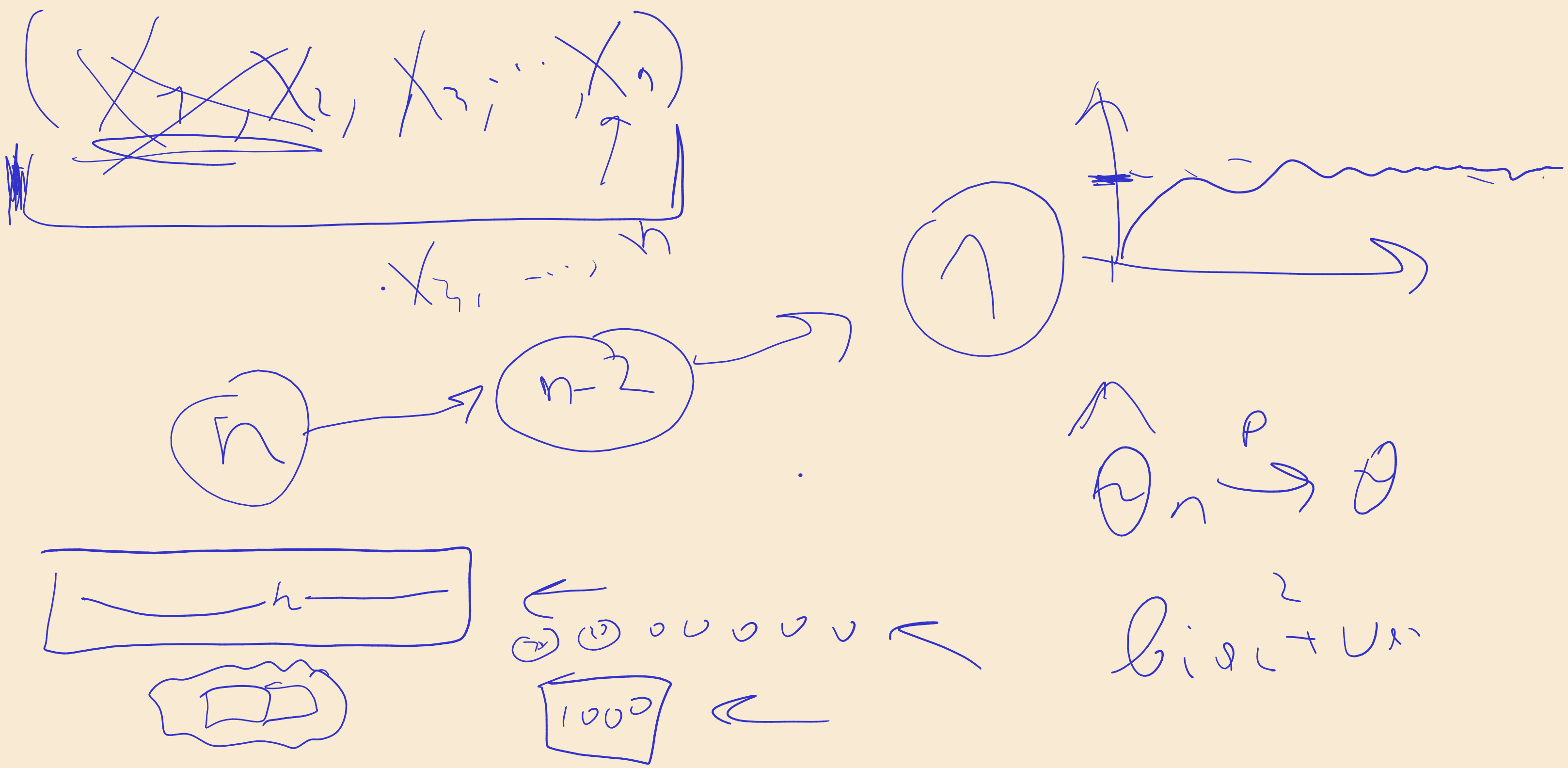
Let $\mathbf{X} = (X_1, \ldots, X_n)^\top$ with $X_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(\pi)$, where $\pi \in (0, 1)$. Define $U(\mathbf{X}) = \sum_{i=1}^n X_i$.

- a. Show that $U(\mathbf{X})/n$ is unbiased for π.
- b. Show that $U(\mathbf{X})$ is **minimal** sufficient for π.
- c. Now consider the estimator for π^2:

$$V(\mathbf{X}) = \frac{U(\mathbf{X})\,[U(\mathbf{X}) - 1]}{n(n-1)}.$$

  Verify that $V(\mathbf{X})$ is unbiased for π^2.

Hint for (c): expand $\mathbb{E}[U(U-1)]$ *using* $\mathbb{E}[U] = n\pi$ *and* $\text{Var}(U) = n\pi(1-\pi)$.



🧀 🧀 03 Normal Variance: Minimal Sufficiency

Let $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ where $\sigma^2 > 0$ is unknown but μ is **known**.

- a. Show that $T(\mathbf{X}) = \sum_{i=1}^{n}(X_i - \mu)^2$ is sufficient for σ^2 using the factorization theorem.
- b. Using the likelihood ratio criterion, show that $T(\mathbf{X})$ is **minimal** sufficient for σ^2.

Recall: T *is minimal sufficient iff* $T(\mathbf{x}) = T(\mathbf{y}) \iff \frac{f(\mathbf{x}|\sigma^2)}{f(\mathbf{y}|\sigma^2)}$ *is free of* σ^2.

$$T(x) = \underbrace{\sum_{i=1}^{n}(x_i - \mu)^2}$$ suff.

$$f(x|\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) = g(T(x), \sigma^2)$$

$T(x)$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \underbrace{\sum_{i=1}^{n}(x_i-\mu)^2}_{T(x)}\right)$$

$\forall x, y$: $\dfrac{f(x \mid \sigma^2)}{f(y \mid \sigma^2)}$ $\sigma^2 (\Leftrightarrow)$

$$\boxed{\exp\left(\frac{1}{2\sigma^2}\left(T(x) - T(y)\right)\right)} \qquad T(x) - T(y)$$

$\sigma^2 = 50^8$

$\sigma^2 = 4$

$\sigma^2 \to \infty$

$\Rightarrow$ $e^0 = 1$

$T(x) = T(y) \to$ minimal