1 Flipping Coins and Hypothesizing

Let Y be a random variable indicating the number of flips until a head is obtained. Then $Y \sim \begin{cases} \text{Geometric}(p), & X = 0 \\ \text{Geometric}(q), & X = 1 \end{cases}$. $L(y) = \frac{\mathbb{P}(Y=y|X=1)}{\mathbb{P}(Y=y|X=0)} = \frac{(1-q)^{y-1}q}{(1-p)^{y-1}p} = \left(\frac{1-q}{1-p}\right)^{y-1}\frac{q}{p}$. Since $1-q < 1-p$, $L(y)$ is decreasing in y, and so $\hat{X} = \begin{cases} 1, & \text{if } Y < y_0 \\ 1, & \text{w.p. } \gamma, \text{ if } Y = y_0 \\ 0, & \text{if } Y > y_0 \end{cases}$. With $\mathbb{P}(\hat{X} = 1 \mid X = 0) = \mathbb{P}(Y < y_0 \mid X = 0) = 1 - \mathbb{P}(Y \geq y_0 \mid X = 0) = 1 - (1-p)^{y_0-1} = \beta$, we get $y_0 = \frac{\ln(1-\beta)}{\ln(1-p)} + 1$. For $Y = y_0$, we then have that $\mathbb{P}(\hat{X} = 1 \mid X = 0) = \gamma\mathbb{P}(Y = y_0 \mid X = 0) = \gamma(1-p)^{y_0-1}p = \beta$, so $\gamma = \frac{\beta}{(1-p)^{y_0-1}p}$.

2 BSC Hypothesis Testing

Let $z_i = x_i \oplus y_i$ where x_i is the ith input and y_i is the ith output. Then, $z_i \sim \text{Bernoulli}(\epsilon)$, so that $Z = \sum_{i=1}^n z_i \sim \text{Binomial}(n, \epsilon)$. We have that $L(Z) = \frac{\mathbb{P}(Z=z|X=1)}{\mathbb{P}(Z=z|X=0)} = \frac{\binom{n}{z}\epsilon'^z(1-\epsilon')^{n-z}}{\binom{n}{z}\epsilon^z(1-\epsilon)^{n-z}} = \left(\frac{\epsilon'(1-\epsilon)}{\epsilon(1-\epsilon')}\right)^z\left(\frac{1-\epsilon'}{1-\epsilon}\right)^n$. $L(Z)$ is increasing in z since $\epsilon'(1-\epsilon) > \epsilon(1-\epsilon')$, so it suffices to just find a threshold z_0. We can use the CLT to approximate a threshold for z_0 such that $\hat{X} = \begin{cases} 1, & \text{if } Z > z_0 \\ 0, & \text{if } Z < z_0 \end{cases}$. We have $.05 \geq \mathbb{P}(\hat{X} = 1 \mid X = 0) = \mathbb{P}(Z > z_0 \mid X = 0) = \mathbb{P}\left(\frac{Z - .1n}{\sqrt{n(.1)(.9)}} > \frac{z_0 - .1n}{\sqrt{n(.1)(.9)}}\right) \approx \mathbb{P}\left(\mathcal{N}(0,1) > \frac{z_0 - .1n}{\sqrt{n(.1)(.9)}}\right) \implies \frac{z_0 - .1n}{\sqrt{n(.1)(.9)}} \approx 1.645$, so that $z_0 \approx 1.645\sqrt{.09n} + .1n$. This optimal decision rule does not depend on the specific choice of ϵ' since $L(Z)$ is always increasing in z for $\epsilon' > \epsilon$.

3 Projections

(a) Since $\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[XY] \leq \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2|\mathbb{E}[XY]| \leq \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\sqrt{\mathbb{E}[X^2] + \mathbb{E}[Y^2]} < \infty$, and $\mathbb{E}[(aX)^2] = a^2\mathbb{E}[X^2] < \infty$, $\mathcal{H}$ is a real vector space as it is closed under addition and scalar multiplication. Moreover, $\mathbb{E}[XY] = \mathbb{E}[YX]$, so that $\langle X, Y\rangle = \langle Y, X\rangle$, and $\mathbb{E}[(X+cY)Z] = \mathbb{E}[(X+cY)Z] = \mathbb{E}[XZ + cZY] = \mathbb{E}[XZ] + c\mathbb{E}[ZY]$, so $\langle X + cY, Z\rangle = \langle X, Z\rangle + c\langle Z, Y\rangle$. $\mathbb{E}[X^2] > 0$ so that $\langle X, X\rangle > 0$. As such $\langle X, Y\rangle \coloneqq \mathbb{E}[XY]$ for $\mathcal{H}$ forms a real inner product space.

(b)

$$\begin{aligned}
||u + cv - x||^2 &= ||u + cv - Tu - cTv + Tu + cTv - x||^2 \\
&= ||u + cv - Tu - cTv||^2 + 2\langle u + cv - Tu - cTv, Tu + cTv - x\rangle + ||Tu + cTv - x||^2 \\
&= ||u + cv - Tu - cTv||^2 + ||Tu + cTv - x||^2 \\
&\geq ||u + cv - Tu - cTv||^2
\end{aligned}$$

(c) Since any finite-dimensional vector space is essentially identical to $\mathbb{R}^n$ and $P = \sum_{i=1}^n v_i v_i^\intercal$ for $\mathbb{R}^n$, we have that $Py = \sum_{i=1}^n v_i v_i^\intercal y = \sum_{i=1}^n \langle y, v_i\rangle v_i$.

4 Exam Difficulties

(a) $\hat{\Theta} = L[\Theta \mid X] = \mathbb{E}[\Theta] + \frac{\text{cov}(\Theta,X)}{\text{var}(X)}(X - \mathbb{E}[\Theta])$, where $\mathbb{E}[\Theta] = 50$, $\text{var}(X) = \text{var}(\mathbb{E}[X \mid \Theta]) + \mathbb{E}[\text{var}(X \mid \Theta)] = \text{var}\left(\frac{\Theta}{2}\right) + \mathbb{E}\left[\frac{\Theta^2}{12}\right] = \frac{1}{4}\frac{100^2}{12} + \frac{1}{12}\int_0^{100} \theta^2 \mathrm{d}\theta = \frac{625}{3} + \frac{100^2}{36} = \frac{4375}{9}$, and $\text{cov}(\Theta, X) = \mathbb{E}[\Theta X] - \mathbb{E}[\Theta]\mathbb{E}[X] = \frac{1}{100}\int_{\theta=0}^{100}\int_{x=0}^{\theta} x \mathrm{d}x \mathrm{d}\theta - 50\mathbb{E}[\mathbb{E}[X \mid \Theta]] = \frac{1}{200}\int_{\theta=0}^{00} \theta^2 \mathrm{d}\theta - \mathbb{E}\left[\frac{\Theta}{2}\right] = \frac{100^3}{600} - 1250 = \frac{1250}{3}$.

(b) $\text{MAP}[\Theta \mid X] = \arg\max_{\Theta} \mathbb{P}(\Theta \mid X) = \arg\max_{\Theta} f_\Theta(\theta) f_{X|\Theta}(x \mid \theta)$. Since Θ is uniform, we need only consider $\arg\max_{\Theta} f_{X|\Theta}(x \mid \theta) = \frac{1}{\theta}$. Since $X \le \Theta \le 100$, we get that $\text{MAP}[\Theta \mid X] = X$.

5 Photodetector LLSE

Let T be the number of transmitted photons. Then $T = \begin{cases} \Theta, & \text{w.p. } p \\ 0, & \text{w.p. } 1-p \end{cases}$.

$L[T \mid T+N] = \mathbb{E}[T] + \frac{\text{cov}(T,T+N)}{\text{var}(T+N)}(T + N - \mathbb{E}[T])$, where $\mathbb{E}[T] = p\mathbb{E}[\Theta] = p\lambda$, $\text{var}(T+N) = \text{var}(T) + \text{var}(N) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 + \mu = \mathbb{E}[\mathbb{E}[T^2 \mid \Theta]] - (p\lambda)^2 + \mu = \mathbb{E}[p\Theta^2] - (p\lambda)^2 + \mu = p(\lambda + \lambda^2) - (p\lambda)^2 + \mu$, and $\text{cov}(T, T+N) = \mathbb{E}[T(T+N)] - \mathbb{E}[T]\mathbb{E}[T+N] = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = \text{var}(T) = p(\lambda + \lambda^2) - (p\lambda)^2$.

6 [Bonus] p-Value