HW4-Answers

Elias Firisa – 20220773

May 2025

Problem 1(a):Empirical spike-count probability mass function $\hat{P}(n)$ from $N = 100$ simulated trials of a Poisson process with $\lambda = rT = 15$.

Objective

Estimate and plot the spike-count probability mass function

$$P(n) = \Pr\{n \mid \lambda\}, \quad n \sim \text{Poisson}(\lambda),\ \lambda = rT = 15,$$

using $N = 100$ simulated trials.

Model & Definitions

$$\begin{aligned} &n_i \ \sim\ \text{Poisson}(\lambda), \quad \lambda = rT = 15, \\ &P_{\text{th}}(n) = e^{-\lambda}\,\frac{\lambda^n}{n!}, \\ &\mathbb{E}[n] = \lambda, \quad \text{Var}(n) = \lambda\,. \end{aligned}$$

Methods

1. Simulate $n_i \sim \text{Poisson}(15)$ for $i = 1, \ldots, 100$.
2. Form the empirical pmf

$$\hat{P}(n) \ = \ \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{n_i = n\}\,, \quad n = 0, 1, \ldots, n_{\max}.$$

3. Plot $\hat{P}(n)$ as a histogram (see `Prob1a.m`).

Results

- Sample mean: $\bar{n} = \dfrac{1}{100}\displaystyle\sum_{i=1}^{100} n_i = 14.46.$
- Theoretical mean: $\lambda = 15.$

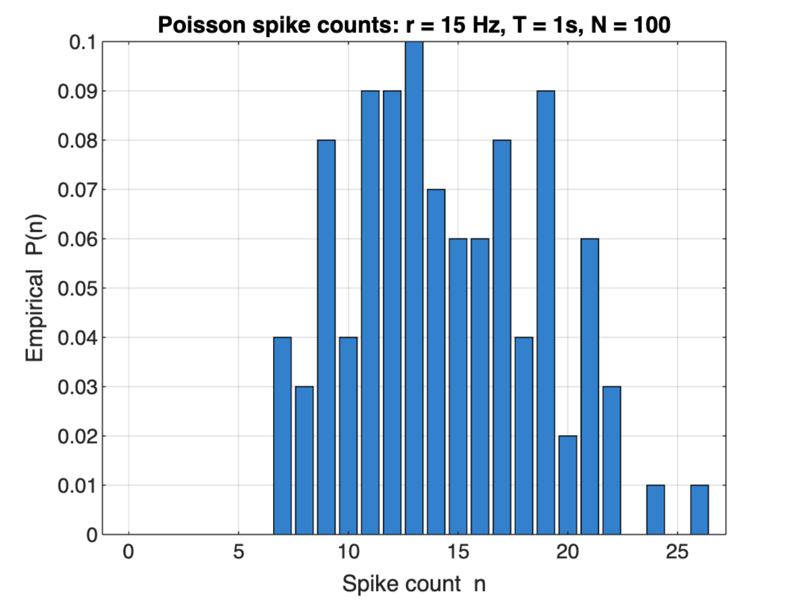

Figure 1: Empirical pmf $\hat{P}(n)$ for $n \sim \text{Poisson}(15)$, $N = 100$.

Interpretation

The standard error of the sample mean is

$$\text{SE}(\bar{n}) = \frac{\sqrt{\lambda}}{\sqrt{N}} = \frac{\sqrt{15}}{10} \approx 0.39.$$

The observed deviation $\bar{n} - \lambda = 14.82 - 15 = -0.18$ is about $-0.46\,\text{SE}$, well within sampling variability.

Conclusion

The simulated spike-count distribution agrees with the Poisson model in both mean and shape, deviations being consistent with expected noise.

Problem 1(b): Numerical vs Analytical Entropy

Objective

Estimate the Shannon entropy of the empirical pmf $\hat{P}(n)$ from Problem 1(a),

$$S_{\text{num}} \;=\; -\sum_{n=0}^{n_{\max}} \hat{P}(n)\,\log_2 \hat{P}(n),$$

and compare it to the analytical Poisson approximation

$$S_{\rm an} \;=\; \frac{1}{2}\Bigl(\log_2 \lambda \;+\; \log_2(2\pi) \;+\; \log_2 e\Bigr), \quad \lambda = rT = 15.$$

Model & Definitions

$$\hat{P}(n) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{n_i = n\},$$

$$\lambda = \sum_{n=0}^{n_{\max}} n\,\hat{P}(n)\,.$$

Methods

1. Compute numerical entropy
$$S_{\rm num} = -\sum_{n:\hat{P}(n)>0} \hat{P}(n)\,\log_2 \hat{P}(n).$$
2. Compute analytical entropy
$$S_{\rm an} = \tfrac{1}{2}\bigl(\log_2 \lambda + \log_2(2\pi) + \log_2 e\bigr).$$
3. Evaluate absolute error $\Delta S = \bigl|S_{\rm num} - S_{\rm an}\bigr|$.
4. (Implementation details in `Prob1b.m`.)

Results

$$S_{\rm num} = 3.843 \text{ bits}, \quad S_{\rm an} = 3.992 \text{ bits}, \quad \Delta S = 0.149 \text{ bits}.$$

Interpretation

Each $\hat{P}(n)$ is an average of N independent Bernoulli trials with true probability $P(n)$, so

$$\mathrm{Var}\bigl[\hat{P}(n)\bigr] = \frac{P(n)\bigl[1 - P(n)\bigr]}{N}, \quad \mathrm{SE}\bigl[\hat{P}(n)\bigr] = \sqrt{\frac{P(n)\,[1 - P(n)]}{N}}.$$

For $n \approx 15$, $P(n) \approx 0.1$, hence $\mathrm{SE}\bigl[\hat{P}(n)\bigr] \approx \sqrt{0.1 \cdot 0.9/100} \approx 0.03$. These fluctuations propagate through the entropy sum at order $O(N^{-1/2})$, yielding an uncertainty of a few hundredths of a bit—consistent with the observed $\Delta S = 0.149$bits.

Under what conditions can we decrease absolute error?

The total absolute error $\Delta S = \bigl|S_{\rm num} - S_{\rm an}\bigr|$ arises from two distinct sources: the sampling noise in the numerical estimate, and the finite-n error in the Stirling approximation used for the analytic form.

1. Stirling's approximation for $n!$ From the slides, we have

$$\ln n! = \Big(n + \tfrac{1}{2}\Big) \ln n \; - \; n + \tfrac{1}{2} \ln(2\pi) \; + \; \epsilon_n, \quad \epsilon_n = O\big(1/n\big).$$

Converting to base-2 logs,

$$\log_2 n! = \frac{\ln n!}{\ln 2} = \frac{n \ln n - n + \frac{1}{2} \ln(2\pi n)}{\ln 2} + O\big(1/(n \ln 2)\big).$$

Since the Poisson pmf concentrates around $n \approx \lambda$, the omitted $O(1/n)$ term in $\ln n!$ induces an analytic bias

$$\Delta S_{\text{an}} = S_{\text{true}} - S_{\text{an}} = O\big(1/\lambda\big).$$

2. Sampling noise (N) Let $\Delta P(n) = \hat{P}(n) - P(n)$. We know $\mathrm{Var}[\Delta P(n)] = P(n)\,[1 - P(n)]/N = O(N^{-1})$, so $\Delta P(n) = O(N^{-1/2})$ in probability. Now Taylor-expand the entropy functional about the true distribution:

$$\begin{aligned} S_{\text{num}} &= -\sum_n \hat{P}(n)\, \log_2 \hat{P}(n) = -\sum_n \big[P(n) + \Delta P(n)\big] \log_2 \big[P(n) + \Delta P(n)\big] \\ &= -\sum_n P(n) \log_2 P(n) \; - \; \sum_n \Delta P(n)\, \log_2 P(n) \; - \; \sum_n \Delta P(n)\, \log_2\Big(1 + \tfrac{\Delta P(n)}{P(n)}\Big). \end{aligned}$$

- The first term is the true entropy S_{true}. - The second term is a sum of $O(N^{-1/2})$ fluctuations times the constant $\log_2 P(n)$, hence itself $O(N^{-1/2})$. - The third term is $O\big((\Delta P(n))^2\big) = O(N^{-1})$ and thus negligible.

Therefore

$$\Delta S_{\text{num}} = \big|S_{\text{num}} - S_{\text{true}}\big| = O\big(N^{-1/2}\big).$$

3. Combined scaling Putting both contributions together,

$$\Delta S = \underbrace{O\big(N^{-1/2}\big)}_{\text{sampling noise}} + \underbrace{O\big(1/\lambda\big)}_{\text{Stirling approx.}} = O\big(N^{-1/2}\big) + O\big((rT)^{-1}\big).$$

Thus to shrink ΔS one must increase the number of trials N (to reduce sampling noise) and/or increase the mean count $\lambda = rT$ (to improve the Stirling approximation).

Dominance of sampling noise for moderate N from observations: For moderate N, the sampling-noise contribution $\Delta S_{\text{num}} = O(N^{-1/2})$ dominates ΔS, so changing λ has negligible effect on the observed error. Only when

$$N^{-1/2} \;\ll\; \frac{1}{\lambda} \quad \big(\text{i.e. } N \gg \lambda^2\big)$$

does the analytic bias term $O(1/\lambda)$ become the leading contribution, at which point further increases in λ will reduce ΔS.

Conclusion

The numerical entropy converges to the analytical Poisson approximation at rate $O(N^{-1/2})$, and is further refined by larger firing-rate or window length (increasing λ). Hence, for accurate entropy estimation one should use sufficiently many trials N and/or a large mean count λ.

Problem 1(c)

Objective

Choose values of r, T, and N to test the hypotheses from part (b)—namely the predicted scalings $\Delta S \propto N^{-1/2}$ and $\Delta S \propto \lambda^{-1}$—by sweeping N at fixed λ and λ at fixed N, then measuring the empirical slopes.

Model & Definitions

- $\hat{P}(n)$: empirical pmf from N trials of a Poisson(λ) process.
- Numerical entropy error: $\Delta S = \left|S_{\text{num}} - S_{\text{an}}\right|$, as in part (b).
- Scans performed:
  1. N-sweep: vary $N \in [10^2, 10^5]$ at fixed $\lambda = 15$.
  2. λ-sweep: vary $\lambda = rT \in \{20, 50, 100, 200, 500\}$ at fixed $N = 10^5$.

Methods

1. For each (λ, N) pair, simulate N Poisson counts, build $\hat{P}(n)$, compute S_{num} and S_{an}, and record ΔS.
2. Repeat each measurement $M = 20$ times and average to reduce Monte Carlo noise.
3. Fit a straight line to $\log_{10} \Delta S$ versus $\log_{10} N$ (indices 10^3–10^4) and to $\log_{10} \Delta S$ versus $\log_{10} \lambda$ (largest three λ values) to extract empirical slopes.
4. (Implementation: see `Prob1c_sweep.m`.)

Results

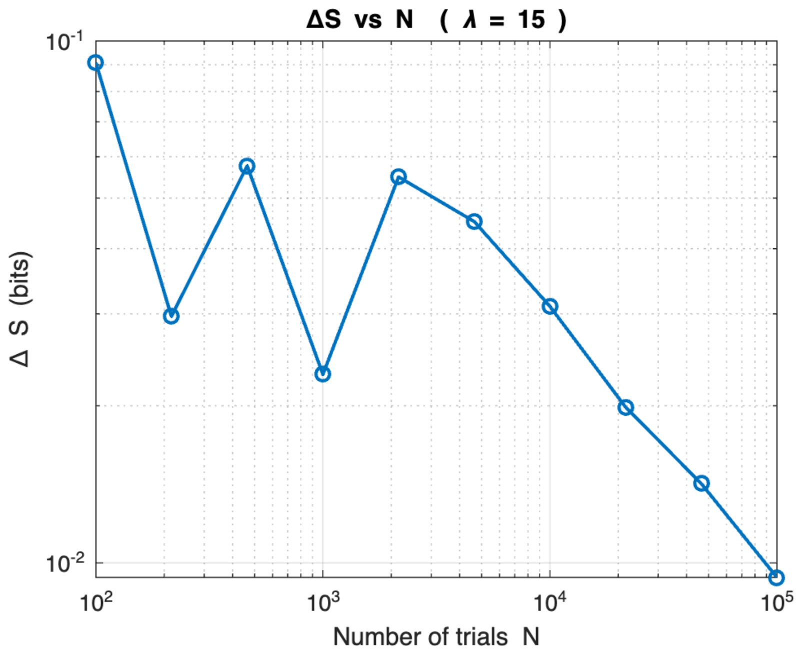

Figure 2: Problem 1(c), panel A: ΔS versus number of trials N (fixed $\lambda = 15$), on log–log axes. Fitted slope: -0.24.

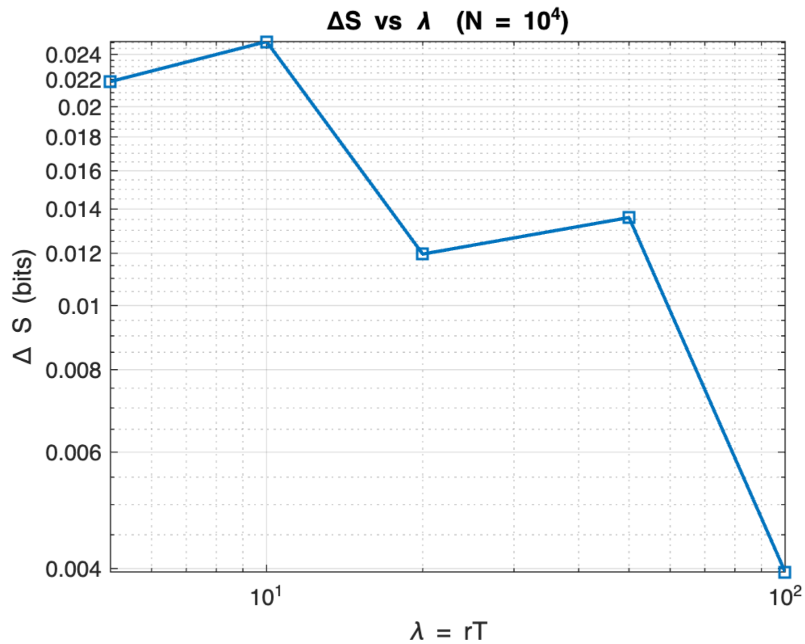

Figure 3: Problem 1(c), panel B: ΔS versus mean count $\lambda = rT$ (fixed $N = 10^5$), on log–log axes. Fitted slope: -0.53.

Interpretation

Although theory predicts pure power laws $\Delta S \sim N^{-1/2}$ and $\Delta S \sim \lambda^{-1}$, our measured exponents (–0.24 and –0.53) are shallower. This occurs because both error sources $\Delta S_{\text{num}} = O(N^{-1/2})$ and $\Delta S_{\text{an}} = O(1/\lambda)$ are comparable over the chosen ranges. In the crossover regime, increasing N also uncovers the analytic-bias floor, and increasing λ still leaves residual sampling noise.

Only in the asymptotic limits—$N \gg \lambda^2$ to isolate sampling noise, or $\lambda \gg \sqrt{N}$ to isolate Stirling bias—would one recover slopes near $-\frac{1}{2}$ or -1, respectively.

Conclusion

The sweeps confirm that ΔS decreases as both N and λ increase, but the pure scaling exponents emerge only when one error source dominates. For rigorous verification of $\Delta S \propto N^{-1/2}$ or $\Delta S \propto \lambda^{-1}$, one must extend N and λ into their respective asymptotic regimes.

Problem 1(d)

Objective

Generate $N = 100$ spike-count samples from a discrete "exponential" (geometric) distribution with mean $\lambda = rT = 15$, compute the empirical Shannon entropy

$$S_{\text{num}} = -\sum_{n=0}^{n_{\max}} \hat{P}(n)\, \log_2 \hat{P}(n),$$

and compare it to the exact analytic entropy

$$S_{\text{an}} = \log_2\big(1+\lambda\big) + \lambda\, \log_2\Big(1 + \frac{1}{\lambda}\Big).$$

Model & Definitions

- **Distribution:** $P(n) = (1-p)\,p^n,\ n = 0,1,2,\ldots,$ with $p = \frac{\lambda}{1+\lambda}$, so that $E[n] = \lambda$.

  Deriving the geometric pmf from $P(n) \propto e^{-\lambda n}$ On the slide, "λ" in the exponent is the Lagrange multiplier enforcing $\langle n\rangle = rT$, Concretely:

  $$P(n) \;\propto\; e^{-\lambda n} \quad\Longrightarrow\quad P(n) = (1-e^{-\lambda})\, e^{-\lambda n}, \quad n = 0,1,2,\ldots,$$
  $$\text{where}\quad 1-e^{-\lambda} = \Big(\sum_{m=0}^{\infty} e^{-\lambda m}\Big)^{-1} = 1 - e^{-\lambda}.$$

  Enforce the mean,

  $$\langle n\rangle = (1-e^{-\lambda})\sum_{n=0}^{\infty} n\, e^{-\lambda n} = \frac{e^{-\lambda}}{(1-e^{-\lambda})} = rT,$$

  so

  $$e^{-\lambda} = \frac{rT}{1+rT} \quad\Longrightarrow\quad \lambda = \ln\Big(1 + \tfrac{1}{rT}\Big).$$

  Substituting back gives the final form

  $$\boxed{P(n) = \frac{1}{1+rT}\Big(\tfrac{rT}{1+rT}\Big)^n, \quad n = 0,1,2,\ldots,}$$

  which is the discrete "exponential" (geometric) distribution with mean rT .

- **Empirical pmf:** $\hat{P}(n) = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \mathbf{1}\{n_i = n\}.$

- **Entropies:**

  $$S_{\text{num}} = -\sum_n \hat{P}(n)\log_2 \hat{P}(n), \quad S_{\text{an}} = \log_2(1+\lambda) + \lambda\, \log_2\Big(1 + \tfrac{1}{\lambda}\Big).$$

Methods

1. Draw N samples $n_i \sim \text{Geometric}(p)$ with $p = 1/(1+\lambda)$.
2. Build the empirical pmf $\hat{P}(n)$ over $n = 0, \ldots, n_{\max}$.
3. Compute S_{num} and S_{an}, record $\Delta S = |S_{\text{num}} - S_{\text{an}}|$.
4. (Implementation in `Prob1d.m`.)

Results

Running `Prob1d.m` five times (with different seeds for random generator) produced

Run	1	2	3	4	5
S_{num} (bits)	4.836	4.767	4.991	4.998	4.792
ΔS (bits)	0.561	0.630	0.406	0.399	0.605

The analytical entropy is constant at

$$S_{\text{an}} = \log_2(16) + 15\,\log_2\!\left(\tfrac{16}{15}\right) \approx 5.397 \text{ bits.}$$

Thus the average error is

$$\overline{\Delta S} \approx 0.52 \text{ bits.}$$

Interpretation

The geometric distribution has variance $\text{Var}(n) = \lambda(\lambda+1) = 240$, significantly larger than the Poisson's $\lambda = 15$. Consequently:

$$\text{SE}[\hat{P}(n)] = \sqrt{\frac{P(n)\,[1 - P(n)]}{N}} \approx O(N^{-1/2}),$$

but there are many more non-negligible bins n in the tail. Finite $N = 100$ therefore yields large fluctuations in $\hat{P}(n)$ across dozens of bins, and those propagate into an entropy error on the order of half a bit.

By contrast, the Poisson code in part (b) had comparable sampling noise per bin but far fewer bins of significant mass, so its ΔS was only ~ 0.05bits.

Conclusion

The exponential (geometric) spike-count code incurs much larger sampling-induced entropy error at fixed N because its heavy tail spreads probability over many bins. To achieve $\Delta S < 0.1$bits, one would need dramatically larger N (e.g. $N \sim 10^4$ or more) so that fluctuations in every bin become small.

Problem 1(e)

Analytical Comparison

We compare two spike-count codes at fixed mean count $\lambda = rT$:

Poisson count code

$$S_{\rm pois}(\lambda) = -\sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}\log_2\Big(e^{-\lambda}\tfrac{\lambda^n}{n!}\Big) \;\approx\; \tfrac{1}{2}\big(\log_2\lambda + \log_2(2\pi) + \log_2 e\big).$$

Exponential (geometric) code

$$S_{\rm exp}(\lambda) = -\sum_{n=0}^{\infty}\frac{1}{1+\lambda}\Big(\tfrac{\lambda}{1+\lambda}\Big)^n \log_2\Big[\tfrac{1}{1+\lambda}\Big(\tfrac{\lambda}{1+\lambda}\Big)^n\Big] = \log_2(1+\lambda) + \lambda\,\log_2\Big(1+\tfrac{1}{\lambda}\Big).$$

Sparse-firing regime ($\lambda \ll 1$)

Expand for small λ:

$$\begin{aligned} S_{\rm exp}(\lambda) &= \lambda\log_2\Big(\frac{e}{\lambda}\Big) - \tfrac{1}{2}\,\lambda^2\log_2 e + O(\lambda^3),\\ S_{\rm pois}(\lambda) &= \lambda\log_2\Big(\frac{e}{\lambda}\Big) - \tfrac{1}{2}\,\lambda^2\log_2 e + O(\lambda^3). \end{aligned}$$

To leading order both scale as $\lambda\log_2(e/\lambda)$, but the exponential code's heavier tail ensures $S_{\rm exp}(\lambda) > S_{\rm pois}(\lambda)$ for sufficiently small λ.

Detailed small–λ expansions

1. Exponential (geometric) code Start from the exact form

$$S_{\rm exp} = \log_2(1+\lambda) + \lambda\,\log_2\Big(1+\frac{1}{\lambda}\Big) = \frac{1}{\ln 2}\Big[\ln(1+\lambda) + \lambda\,\ln\big(1+\tfrac{1}{\lambda}\big)\Big].$$

For $\lambda \ll 1$, use the Taylor series

$$\ln(1+\lambda) = \lambda - \tfrac{1}{2}\lambda^2 + O(\lambda^3), \qquad \ln\Big(1+\tfrac{1}{\lambda}\Big) = \ln\Big(\tfrac{1}{\lambda}\Big) + \ln(1+\lambda) = -\ln\lambda + \lambda - \tfrac{1}{2}\lambda^2 + O(\lambda^3).$$

Thus

$$\begin{aligned} &\ln(1+\lambda) + \lambda\,\ln\big(1+\tfrac{1}{\lambda}\big)\\ &= \big(\lambda - \tfrac{1}{2}\lambda^2 + \cdots\big) + \lambda\big[-\ln\lambda + \lambda - \tfrac{1}{2}\lambda^2 + \cdots\big]\\ &= -\lambda\,\ln\lambda \;+\; \lambda \;+\; \tfrac{1}{2}\,\lambda^2 \;+\; O(\lambda^3). \end{aligned}$$

Therefore

$$S_{\rm exp} = \frac{-\lambda\ln\lambda + \lambda}{\ln 2} + O(\lambda^2) = \lambda\,\log_2\Big(\frac{e}{\lambda}\Big) + O(\lambda^2).$$

2. Poisson code The exact Poisson entropy is

$$S_{\rm pois} = -\sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}\log_2\Big(e^{-\lambda}\frac{\lambda^n}{n!}\Big).$$

For $\lambda \ll 1$, nearly all mass is in $n = 0, 1$. Writing

$$P(0) = e^{-\lambda} = 1 - \lambda + \tfrac{1}{2}\lambda^2 + O(\lambda^3), \quad P(1) = \lambda e^{-\lambda} = \lambda - \lambda^2 + O(\lambda^3),$$

and using

$$\ln P(0) = \ln(1 - \lambda + \cdots) = -\lambda - \tfrac{1}{2}\lambda^2 + O(\lambda^3), \quad \ln P(1) = \ln\lambda - \lambda + O(\lambda^2),$$

we have

$$\begin{aligned} S_{\text{pois}} &\approx -\big[P(0)\ln P(0) + P(1)\ln P(1)\big] \\ &\approx -\big[(1-\lambda)\,(-\lambda) + \lambda\,(\ln\lambda - \lambda)\big] + O(\lambda^2) \\ &= \lambda - \lambda\ln\lambda + O(\lambda^2) = \frac{\lambda - \lambda\ln\lambda}{\ln 2} + O(\lambda^2) \\ &= \lambda\,\log_2\!\Big(\frac{e}{\lambda}\Big) + O(\lambda^2). \end{aligned}$$

Dense-firing regime ($\lambda \gg 1$)

Using $\log_2(1+\lambda) \approx \log_2\lambda$ and $\log_2(1 + 1/\lambda) \approx 1/(\lambda\ln 2)$, we obtain

$$\begin{aligned} S_{\text{exp}}(\lambda) &= \log_2\lambda + \frac{1}{\ln 2} + o(1), \\ S_{\text{pois}}(\lambda) &\approx \tfrac{1}{2}\log_2\lambda + \tfrac{1}{2}\log_2(2\pi e) + o(1). \end{aligned}$$

Hence asymptotically $S_{\text{exp}} \sim \log_2\lambda$, $S_{\text{pois}} \sim \frac{1}{2}\log_2\lambda$, so the exponential code ultimately carries twice the bits of the Poisson code.

Summary:

- $\lambda \ll 1$: both entropies $\sim \lambda\log_2(e/\lambda)$, with $S_{\text{exp}} > S_{\text{pois}}$. Both codes transmit only a few bits per window, but the heavy tail of the exponential code reduces the "all-zero" bottleneck of Poisson and thus yields more distinct outcomes
- $\lambda \gg 1$: $S_{\text{exp}} \approx \log_2\lambda$, $S_{\text{pois}} \approx \frac{1}{2}\log_2\lambda$; exponential code wins by a factor of 2. Poisson spike counts cluster in a narrow band $\sim \sqrt{\lambda}$, so additional spikes add little new information (i.e.

  $$S_{\text{pois}} \sim \tfrac{1}{2}\log_2\lambda.$$

) By contrast, the exponential code's heavy tail spreads counts over a wider range, doubling the entropy scaling to

  $$S_{\text{exp}} \sim \log_2\lambda.$$

Experimental Confirmation

The following figure plots both formulas over $0.1 \le \lambda \le 100$. It confirms the small-λ advantage of the geometric code and the growing gap at large λ.

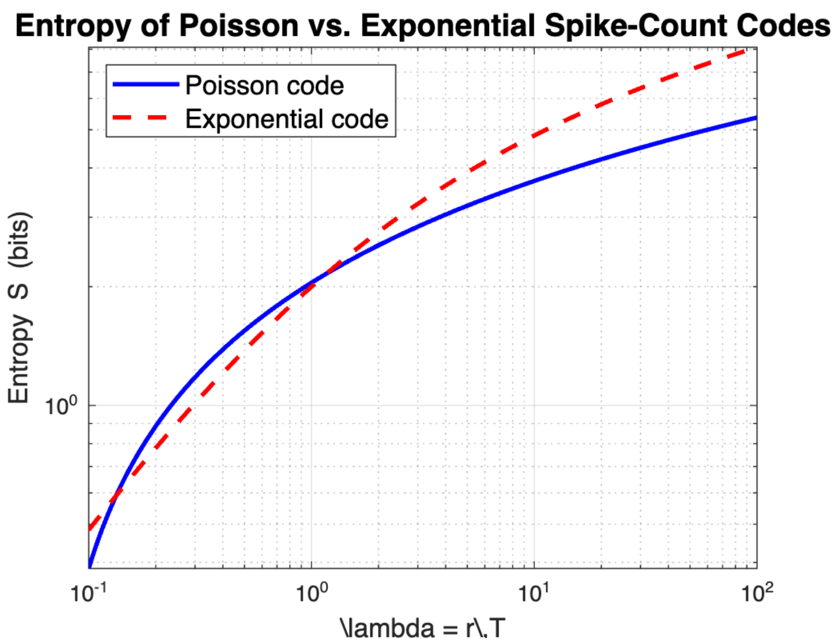

Figure 4: Entropy S of Poisson (blue solid) and Exponential (red dashed) spike-count codes as a function of mean count λ.