A Simple Bayesian Model of Hiring Signals via Top-Coauthor and Publication Count

January 20, 2026

1 Timing, information, and payoffs

We consider a simple two-step process:

1. **Application.** A candidate applies for the job.

2. **Hiring decision.** The employer observes signals about the candidate and decides whether to hire.

Key variables and parameters.

- $T \in \{G, B\}$: candidate type (good G or bad B), with $\Pr(T = G) = p$.

- $A \in \{0, 1\}$: indicator for having published/worked with a high-rank (top) researcher ($A = 1$ yes, $A = 0$ no).

- $N \in \{0, 1, 2, \dots\}$: general quality signal given by the number of papers.

- $w > 0$: applicant payoff if hired.

- B (if $T = G$) and b (if $T = B$): employer payoff from hiring; typically $B > b$.

The employer observes two pieces of information:

- A **general signal of quality** given by the publication count $N \in \{0, 1, 2, \dots\}$ (the number of papers).

- Whether the candidate **published/worked with a high-rank (top) researcher**, represented by $A \in \{0, 1\}$.

Payoffs are as follows:

- The applicant receives a positive payoff $w > 0$ if hired (and 0 otherwise).

- The employer receives B if the hired candidate is good, and b if the hired candidate is bad (typically $B > b$).

Given the observed signals (paper count and whether there is a top coauthor), the employer forms a posterior belief about the probability that the candidate is good. Call this probability P_{good}.

If the employer hires, their expected (future) payoff is simply:

$$\text{Expected employer payoff} = B \cdot P_{\text{good}} + b \cdot (1 - P_{\text{good}}) = b + (B - b) P_{\text{good}}.$$

2 Setup

We now formalize the probabilistic model. The latent type $T \in \{G, B\}$ and the observed signals (A, N) are as defined in Section 1, with prior $\Pr(T = G) = p$.

3 Likelihood model

3.1 Top-researcher signal

Let
$$\Pr(A = 1 \mid T = G) = q_a, \qquad \Pr(A = 1 \mid T = B) = q_b,$$

with $q_a > q_b$. Equivalently,
$$\Pr(A = 0 \mid T = G) = 1 - q_a, \qquad \Pr(A = 0 \mid T = B) = 1 - q_b.$$

3.2 Publication count (Poisson)

We model the paper count using Poisson distributions:
$$N \mid (T = G) \sim \text{Poisson}(\lambda_a), \qquad N \mid (T = B) \sim \text{Poisson}(\lambda_b).$$

Thus for $n \in \{0, 1, 2, \dots\}$,
$$\Pr(N = n \mid T = G) = e^{-\lambda_a} \frac{\lambda_a^n}{n!}, \qquad \Pr(N = n \mid T = B) = e^{-\lambda_b} \frac{\lambda_b^n}{n!}.$$

3.3 Conditional independence

We assume that, conditional on type, the two signals are independent:
$$(A \perp N) \mid T.$$

Hence
$$\Pr(A, N = n \mid T) = \Pr(A \mid T) \Pr(N = n \mid T).$$

4 Posterior probability of being good

Given an observation $(A = a, N = n)$, Bayes' rule gives

$$\Pr(T = G \mid A = a, N = n) = \frac{\Pr(A = a, N = n \mid T = G) \Pr(T = G)}{\Pr(A = a, N = n \mid T = G) \Pr(T = G) + \Pr(A = a, N = n \mid T = B) \Pr(T = B)}$$

$$= \frac{p \, \Pr(A = a \mid G) \, \Pr(N = n \mid G)}{p \, \Pr(A = a \mid G) \, \Pr(N = n \mid G) + (1 - p) \, \Pr(A = a \mid B) \, \Pr(N = n \mid B)}.$$

With the Poisson likelihoods, the factor $n!$ cancels, yielding the closed form

$$\boxed{\Pr(T = G \mid A = a, N = n) = \frac{p \, \Pr(A = a \mid G) \, e^{-\lambda_a} \lambda_a^n}{p \, \Pr(A = a \mid G) \, e^{-\lambda_a} \lambda_a^n + (1 - p) \, \Pr(A = a \mid B) \, e^{-\lambda_b} \lambda_b^n}.}$$

For convenience, define

$$\Pr(A = a \mid G) = \begin{cases} q_a & a = 1 \\ 1 - q_a & a = 0 \end{cases} \qquad \Pr(A = a \mid B) = \begin{cases} q_b & a = 1 \\ 1 - q_b & a = 0 \end{cases}.$$