Last time:

1. Encoding & decoding
2. Reverse correlation
3. GLMs
4. Stimulus/noise correlations & pop. coding

\[ p_{sc} \]

\[ SNR \propto \frac{N^2}{N + N(N-1)p_{sc}} \]

Identical tuning

\[ r_1 = \Theta s + \eta_{common} + b \]

\[ r_2 = -\Theta s + \eta_{common} + b \]

\[ R = r_1 - r_2 = 2\Theta s \]
\( \rho (c_1, z_1) \)

\( \rho (c_1, z_2) \)

\( \rho_{\text{stim}} > 0 \)
\( \rho_{\text{sc}} = 0 \)

\( \rho_{\text{stim}} > 0 \)
\( \rho_{\text{sc}} > 0 \)

\( \rho_{\text{stim}} < 0 \)
\( \rho_{\text{sc}} > 0 \)

"w/o knowledge of cues"

"optimal"
"Sign rule: If sign of $p_{sc}$ & $p_{stim}$ are opposite
forall pairs, Fisher information (& other metrics) are
greater than trial-shuffled information.

Population decoding: Example

$$
\lambda_i = r_{max} \exp \left( -\frac{(s - s_i)^2}{2\sigma^2} \right)
$$

$\lambda_i$ preferred stimulus
tiny width

$$
r_i \sim \text{Pois}(\lambda_i)
$$

If independent,

$$
P(r|s) = \prod_i P(r_i|s)
$$
\[ P(r|s) = \prod_i \frac{\lambda_i^{r_i}}{r_i!} e^{-\lambda_i} \]

\[ \log P(r|s) = \sum_i r_i \log \lambda_i(s) - \log r_i! - \lambda_i(s) \]

Simplifying assumption: \( \sum_i \lambda_i(s) \approx \text{const.} \)

Maximize \( \sum_i r_i \log \lambda_i(s) = \sum_i r_i \frac{-(s-s_i)^2}{2\sigma_i^2} + C \)

Derivative wrt \( s \) \( = 0 \) \( \Rightarrow \)

\( 0 = \sum_i r_i \frac{-2(s-s_i)}{2\sigma_i^2} \)

\( s = \frac{\sum_i r_i s_i / \sigma_i^2}{\sum_i r_i / \sigma_i^2} \)

Preferred stimulus of neuron \( i \) weighted by \( \frac{r_i}{\sigma_i^2} \)

If \( \sigma_i \) constant, reduces to weighted sum

\( s = \frac{\sum_i r_i s_i}{\sum_i r_i} \)}
Maximum likelihood estimate.

Bayes rule: \( P(s|r) = \frac{P(r|s)P(s)}{P(r)} \)

Maximum a posteriori estimate: Maximize \( P(s|r) \) given prior \( P(s) \) \( \iff \)

Maximize \( P(r|s)P(s) \). If \( P(s) = \text{constant} \) (flat prior), MAP estimate = MLE

Example: Gaussian \( P(s) = \frac{1}{\sqrt{2\pi} \sigma_{\text{prior}}^2} \exp\left(-\frac{(s-s_{\text{prior}})^2}{2\sigma_{\text{prior}}^2}\right) \)

MAP: \( \max \log P(r|s)P(s) \)

\[ = \sum_i r_i \cdot \frac{-(s-s_i)^2}{2\sigma_i^2} - \frac{(s-s_{\text{prior}})^2}{2\sigma_{\text{prior}}^2} \]

\[ \implies s_{\text{MAP}} = \frac{\sum_i r_i s_i / \sigma_i^2 + s_{\text{prior}} / \sigma_{\text{prior}}^2}{\sum_i r_i / \sigma_i^2 + s_{\text{prior}} / \sigma_{\text{prior}}^2} \]

What about correlated case? Can no longer factorize likelihood.
Bias & variance:

Bias of $\hat{s}$: $b(s) = E[\hat{s} | s] - s$

(defined for any parameter $\Theta$ of a statistical model)

$\text{Var}(\hat{s} | s) = E\left[ (\hat{s} - E[\hat{s} | s])^2 | s \right]$

$E[(\hat{s} - s)^2] = E\left[ (\hat{s} - E[\hat{s} | s] + b(s))^2 | s \right]$

$= \text{Var}(\hat{s} | s) + b^2(s)$

---

$\text{bias} = \frac{\text{variance}}{\text{bias}}$

$b(s) = 0$ : unbiased estimator

Fisher information (scalar case):

$$I_F(s) = -E\left[ \frac{\frac{d^2}{ds^2} \log p(x | s)}{ds^2} \right]_{p(x | s)}$$  \hspace{1cm} (1)

$$= -\int dx \cdot p(x | s) \frac{d^2}{ds^2} \log p(x | s) \left[ \frac{1}{ds} \right]^2 \frac{1}{E\left[ (\frac{d}{ds} \log p(x | s))^2 \right]}_{p(x | s)}$$  \hspace{1cm} (2)
Example: \( \lambda_i \sim \mu(s) \) \( r_i \sim \text{Pois}(\lambda_i) \)

\[
p(i; s) = \frac{\mu(s)^{r_i} e^{-\mu(s)}}{r_i !}
\]

\[
\log p(v; s) = r_i \log \mu(s) - \mu(s) - \log r_i !
\]

\[
\frac{d}{ds} \log p(v; s) = r_i \frac{\mu'(s)}{\mu(s)} - \mu'(s) = \mu'(s) \left[ \frac{r_i}{\mu(s)} - 1 \right]
\]

\[
E[v^2; s] = \left( \frac{\mu'(s)}{\mu(s)} \right)^2 E \left[ \frac{r_i^2}{\mu(s)} - \frac{2r_i}{\mu(s)} + 1 \right] \quad E[r_i] = \mu(s)
\]

\[
E[i^2] = \frac{1}{\mu(s)} + 1 - 2 + 1
\]

\[
\frac{(\mu'(s))^2}{\mu(s)}
\]
\[ \lambda_i = \max \exp \left( \frac{-(s_i - s)^2}{2\sigma_i^2} \right) \quad i \sim \text{Pois} (\lambda_i) \]

\[ I_F(s) = \frac{\max (s_i - s)^2}{\sigma_i^4} \exp \left( \frac{-(s_i - s)^2}{2\sigma_i^2} \right) \]

**Intuition (using def 1):** Expected curvature of log-likelihood fn.

**Intuition (using def. 2):** "Score": \( \frac{d}{ds} \log p(x|s) \)

How much does log-likelihood of observing \( x \) change when \( s \) is varied?

\[
E \left[ \frac{d}{ds} \log p(x|s) \right] = \int dr \ p(x|s) \cdot \frac{d}{ds} \frac{p(x|s)}{p(x|s)}
\]

\[ = \frac{d}{ds} \int dr \ p(x|s) = \frac{d}{ds} (1) = 0. \]

So \( \text{Var(sore)} = E [\text{score}^2] = I_F(s) \)
**Properties:**

1. Local (dependent on value of $s$)

2. If $r_i$, $r_j$ independent, $I^i_F(s) = I^{i, i}_F(s) + I^{i, j}_F(s)$

   $\log p(r | s) = \log p(r | s) + \log p(r | s)$

3. More generally, $I^{x, y}_F = I^{x}_F + I^{y | x}_F$

4. Dependent on stimulus parameterization:

   If $u = g(s)$, $I^i_F(s) = I^i_F(u) \left( \frac{du}{ds} \right)^2$


**Equivalence of 2 definitions:**

\[
\frac{d^2}{ds^2} \log p(r | s) = \frac{d}{ds} \left[ \frac{d}{ds} \log p(r | s) \right]
\]

\[
= \frac{d}{ds} \left[ \frac{1}{p(r | s)} \frac{d}{ds} p(r | s) \right]
\]

\[
= \frac{1}{p(r | s)} \frac{d^2}{ds^2} p(r | s) - \left( \frac{\frac{d}{ds} p(r | s)}{p(r | s)} \right)^2
\]

\[
E[\cdot] = \int dr \frac{d^2}{ds^2} p(r | s)
\]

\[
= \frac{d^2}{ds^2} \int dr p(r | s) = 0
\]

So

\[
I_F(s) = E \left[ \left( \frac{d}{ds} \log p(r | s) \right)^2 \right] p(r | s)
\] (2)