In all your derivations please show your work in conplete detail. Cite explicitely all theorems and standard results you assume known. Highlight all assumptions needed for your conclusions to be valid.

Problem 1: Consider $X_1, X_2, \ldots, X_n$ iid Bernoulli(p). Under squared error loss, the Bayesian estimator for p is the posterior mean

$$\gamma_n = \frac{a + \sum_i X_i}{a + b + n}$$

(a) Find the MLE of p and characterize its limiting distribution.
(b) Is γ_n consistent for p?
(c) Is γ_n an asymptotically efficient estimator of p?

Problem 2: Consider $X_1, X_2, \ldots, X_n$ iid $N(\mu \neq 0, \sigma^2 < \infty)$.
Let $S_n^2 = \frac{1}{n}\sum_i (X_i - \bar{X}_n)^2$. Characterize the limiting distribution of $S/\bar{X}_n$

Problem 3: Let $Y_1, Y_2, \ldots, Y_n$ be independent Bernoulli(p_i) random variables, with $\log\left(\frac{p_i}{1-p_i}\right) = x_i^T \beta$. In the foregoing formulation, $x_i \in \mathbb{R}^k$, $(i = 1, \ldots, n)$ are k covariate measurements.

(a) Find the explicit forms for $E(Y_i)$ and $Var(Y_i)$.
(b) Let $\mu_i(\beta) = E(Y_i)$. Furthermore, let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^T$, $\mu(\beta) = (\mu_1, \mu_2, \ldots, \mu_n)^T$, and $\mathbf{X} \in \mathbb{R}^{n \times k}$, be a matrix of covariate measurements. Show that the score function $\ell_n'(\beta)$ can be written in the form

$$\ell_n'(\beta) = \mathbf{X}^T[\mathbf{Y} - \mu(\beta)].$$

Problem 4: Consider two stochastic sequences $\{X_n\}_{n \in \mathbb{N}}$, and $\{Y_n\}_{n \in \mathbb{N}}$, such that

$$X_n \to_d X, \text{ and } Y_n \to_p c,$$

where X is a random quantity and c is a constant. In this context, Slutsky's theorem states that:

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \to_d \begin{pmatrix} X \\ c \end{pmatrix}.$$

Let Y be a non-degenerate random quantity. Construct a counterexample to show that Slutsky's result may not be strengthened by changing $Y_n \to_p c$ to $Y_n \to_p Y$.

Problem 1

NAME: David Levy

Convergence in distribution is weaker than convergence in probability, so this seems strange. But we want to construct an example where:

$$\vec{X}_n = (X_{n1}, X_{n2}) \text{ and } \vec{X} = (X_1, X_2)$$

We want

$X_{n1} \xrightarrow{P} X_1$ and $X_{n2} \xrightarrow{P} X_2$, so that we are guaranteed $f(\vec{X}_n) \xrightarrow{P} f(\vec{X})$. We don't need to prove that. We do need to show that $X_{n1} + X_{n2} \stackrel{d}{\not\to} X_1 + X_2$. If we choose $X_{n1} = Z_1 \sim n(0,1)$ and $X_{n2} = Z_2 \sim n(1,1)$, then we know that $X_{n1} \xrightarrow{P} 0$ and $X_{n2} \xrightarrow{P} 1$, but we can see that $X_{n1} + X_{n2} \xrightarrow{d} n(1,2)$ which shows that $f(\vec{X}_n) \stackrel{d}{\not\to} f(\vec{X})$.

- Note that it is also true that $X_{n1} \xrightarrow{P} Z \sim N(0,1)$ and $X_{n2} \xrightarrow{P} W \sim N(1,1)$.
- The statement $X_{n1} + X_{n2} \xrightarrow{d} N(1,2)$ is correct and so is $X_{n1} + X_{n2} \xrightarrow{P} Z + W$.
- This would have worked if $\mathrm{Cov}(X_{n1}, X_{n2}) \neq 0 \ \forall n$.

1	2/5
2	5/10
3	10/10
4	2/5
	19/30

Problem 2

$X_i \sim \text{Poisson}(\lambda) \quad f(x|\lambda) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$

a) I think we need to get a mle $\hat{\lambda}$ of λ, and since we are dealing with an exponential family, that will allow us to use $\hat{\lambda}$ for a UMP size α test.

$L(\lambda; x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$

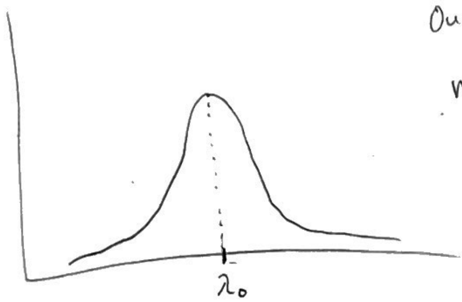
↳ We know this should turn out to be $\bar{X}$ because λ is the mean and variance

$\ell(\lambda; x_1, \ldots, x_n) = \sum_{i=1}^{n} \left[-\lambda + x_i \log\lambda - \log(x_i!)\right]$

$= -n\lambda + \log\lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log(x_i!)$

$\ell'(\lambda) = -n + \left(\frac{1}{\lambda}\right) \sum_{i=1}^{n} x_i \implies \hat{\lambda} = \bar{X}$ is the mle

We should point out that we can use likelihood ratios to help construct the UMP, where we focus on the boundary $\lambda = \lambda_0$



Our statistic has a $N(\lambda_0, \lambda_0/n)$ distribution under the null hypothesis

We can get a useful result from the exponential form, so

$f(x|\lambda) = \exp\left[\underbrace{x_i}_{K(x_i)} \underbrace{\log\lambda}_{p(\lambda)} - \underbrace{\log(x_i!)}_{s(x)} \underbrace{-\lambda}_{q(\lambda)}\right]$

$p(\theta_1 - \theta_2) \sum_{i=1}^{n} k(x_i) - q [$

c) $\pi(\lambda) \propto \lambda^{\alpha-1} e^{-b\lambda}$

prior

joint pdf

$$\pi(\lambda \mid x_1, \dots, x_n) = \frac{\pi(\lambda) f(x_1, \dots, x_n \mid \lambda)}{m(x_1, \dots, x_n)}$$

posterior

marginal pdf

We know $f(x_1, \dots, x_2 \mid \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$, so

$$\pi(\lambda) f(x_1, \dots, x_n \mid \lambda) \propto \left(\lambda^{\alpha-1} e^{-b\lambda}\right) \frac{e^{-n\lambda} \lambda^{\Sigma_i x_i}}{\prod_{i=1}^{n} (x_i!)}$$

$$= \frac{e^{\lambda[-(n+b)]} \lambda^{\alpha-1+\Sigma_i x_i}}{\prod_{i=1}^{n} (x_i!)}$$

Now we need the marginal distribution (likelihood of obtaining our data), which we can get by "integrating out" λ from the joint pdf, which really means

Will also be gamma, because gamma is conjugate to poisson. ✓

d) $E(\lambda \mid x_1, \ldots, x_n) = \sum_{i=0}^{n} \lambda \frac{e^{-\lambda} \lambda^{-x_i}}{}$

↳ limiting will be Normal$(\lambda, \lambda/n)$ ×

Problem 3

b) $L(\alpha; x_1, \ldots x_n) = \prod_{i=1}^{n} \frac{x_i^{\alpha-1} e^{-x_i}}{\Gamma(\alpha)}$ is the likelihood function

$\ell(\alpha; x_1, \ldots, x_n) = \sum_{i=1}^{n} \log\left[\frac{x_i^{\alpha-1} e^{-x_i}}{\Gamma(\alpha)}\right] = \sum_{i=1}^{n}\left[(\alpha-1)\log x_i - x_i - \log\Gamma(\alpha)\right]$ is log likelihood

$$= (\alpha-1)\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n} x_i - n\log\Gamma(\alpha)$$

$$\ell'(\alpha; x_1, \ldots, x_n) = \sum_{i=1}^{n}\log x_i - (n)\left(\frac{1}{\Gamma(\alpha)}\right)\Gamma'(\alpha)$$

$$= \sum_{i=1}^{n}\log x_i - n\psi(\alpha) \quad ✓$$

Setting equal to 0, we get

$$\hat{\alpha} = \psi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i\right) \quad ✓$$

Of course, we need to verify that $\ell''(\alpha) < 0$ for a maximum, and check the bounds to be sure we have found a global maximum.

$\ell''(\alpha) = -n\psi'(\alpha)$ which must be negative by the assumption stated in the problem.

At the bounds we will assume that there is no value exceeding the likelihood evaluated at $\hat{\alpha}$.

We know mles are consistent, so $\hat{\alpha}$ is consistent.

(problem 3)

a) Very simply, the method of moments gives us:

$E X = \alpha/b = \alpha$ and $m_1 = \frac{1}{n} \sum_{i=1}^{n} X = \bar{X}$ ✓

Consistency can be shown by observing that a sum of gamma r.v.s has a distribution as indicated in the hint. So if we define $Y = \sum_{i=1}^{n} X_i$ then $Y \sim \text{Gamma}(n\alpha, 1)$. Thus, $E(Y) = n\alpha/1 = n\alpha$.

Now, $E(\bar{X}) = \frac{1}{n} E(\sum_{i=1}^{n} X) = \frac{1}{n} E(Y) = \alpha$. Thus, $\tilde{\alpha} = \bar{X}$ is an unbiased and consistent estimator of α.

$\tilde{\alpha} \xrightarrow{P} \alpha$ can also be shown using the definition of convergence in probability, $\lim_{n\to\infty} P(|\tilde{\alpha} - \alpha| < \varepsilon) = 1 \quad \forall \varepsilon > 0$, and Chebyshev's inequality, which ultimately gives us the desired result. ✓

(problem 3)

c) We know by CLT that the variance of $\bar{X} = \tilde{\alpha}$ should be $(\alpha/b^2)/n \doteq \alpha/n$, which we can approximate with $\bar{X}/n$.

We can obtain the Fisher information for our mle $\hat{\alpha}$ with $I(\hat{\alpha}) = -E(\ell''(\alpha; x_1, \ldots x_n))$ evaluated at $\hat{\alpha}$.

We know $\ell''(\alpha) = -n\psi'(\alpha)$, and so $-E[-n\psi'(\hat{\alpha})] = nE\psi'(\hat{\alpha}) = n\left(\frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})}\right)$

So $\text{Var}(\hat{\alpha})$ should be $\frac{1}{n}\left(\frac{\Gamma(\hat{\alpha})}{\Gamma'(\hat{\alpha})}\right)$, which will attain a Rao-Cramér lower bound, making $\hat{\alpha}$ the lowest variance estimator (unbiased).

We always want the estimator with the lowest variance, so we prefer $\hat{\alpha}$. ✓

Problem 4

a) $L_2(\theta; x_1, x_2) \propto \exp[-(\bar{x}-\theta)^2] = \exp[-\bar{x}^2 + 2\bar{x}\theta - \theta^2]$

I was on my way,
but erased

b) $L_2(\theta; x_2, x_1) = \left[(2\pi)^{-1/2} \exp\left[-\frac{1}{2}(x_1-\theta)^2\right]\right]\left[(2\pi)^{-1/2} \exp\left[-1/2(x_2-\mu)^2\right]\right]$

$= (2\pi)^{-1} \exp\left[-1/2 \sum_{i=1}^{2}(x_i-\theta)^2\right]$

$\log L_2(\theta; x_2, x_1) = \log[(2\pi)^{-1}] - \frac{1}{2}\sum_{i=1}^{2}(x_i-\theta)^2$

$\log L_2'(\theta) =$

you would finish getting derivative, set equal to zero, then check double derivative for negativity to ensure maximum. you'd check endpoints too.

mle distributed $N(\theta, \frac{1}{n})$ so 95%

CI based on $(1.96)(\frac{2}{n})$ SE.

c) You would use $\pi(\theta | x_1, x_2) = \dfrac{\overset{\text{prior}}{\pi(\theta)}\ \overset{\text{joint}}{f(x_1, x_2|\theta)}}{\underset{\text{marginal}}{m(x_1, x_2)}}$

d) No; you wouldn't. To justify mathematically you would need to show that the marginal and joint yield a different posterior. X

e) No, it won't be the same because the distribution of this statistic will be different. ✓