Probabilistic inference for data science 2, Exam 17.4.2020

Answer to all questions. Calculators are allowed. Computer is allowed to do computations and coding (add your own code to your answer), but no Googling! No copying answers from the friend (if I notice that two answers are exactly the same I will mark 0 points to both!

1. Let $X_1, \dots, X_n$ be an iid sample from $\text{Gamma}(\alpha, \beta)$. Assume α is known. And Gamma pdf is:

$$\text{Gamma}(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\{-x/\beta\} \quad 0 \leq x \leq \infty, \quad \beta > 0 \tag{1}$$

population. Find the MLE of β.

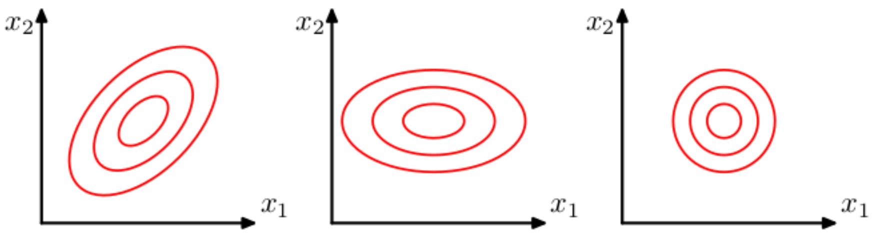
2. Let $X_i = X(u_i)$ be random variable X at spatial locations u_1, u_2 and u_3. Let X_i be normally distributed with common mean μ, common variance σ^2 and correlation $\rho(X_i, X_j) = e^{-\|u_i - u_j\|/2}$. Consider locations $u_1 = (1, 0)$, $u_2 = (0, 0)$ and $u_3 = (0, 3)$.
   (a) Present the distribution of $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$
   (b) Present the likelihood of sample $\boldsymbol{X} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$.

3. Present the Central Limit Theorem for an iid population with known variance. Explain how does it generalize to populations with unknown variance.

4. Are the following statements true or false? Explain your answer, writing **just** true or false, will give you 0 points.
   (a) Let $\mathbf{x}$ be a bivariate Gaussian random variable, i.e. $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The $\boldsymbol{\Sigma}$ is same in cases a), b) and c) in the Figure below?
   (b) Sampling from the infinite population is the same as sampling from a finite population.
   (c) Given a set of scalars $\{x_1, x_2, \dots, x_N\}$, if you want to represent these N scalars by just one scalar, then the sample mean is the best approximator in terms of squared error.
   (d) You will always want your estimator to be unbiased.
   (e) If $\mathsf{Var}X = \sigma^2$, then $P(|X - \mathsf{E}[X]| \geq 2\sigma) \leq \frac{1}{4}$. Additional question: what does the result mean?

5. Bootstrapping, use computer and add code and answer as your solution.
   (a) Use nonparametric bootstrap to estimate the standard deviation of the mean of the data below. Use M=10000 samples. Report sample mean, mean of the bootstrapped means and standard deviation of the bootstrapped means. How does sample mean compare to the mean of bootstrapped means?
   (b) There are two outliers in the data: 6th and 11th items. Remove them, and repeat the above procedure. Report same statistics as above. Compare the standard deviation obtained with and without the outliers.

   Data: 576, 580, 653, 635, 555, 875, 558, 661, 545, 578, 651, 972, 666, 605, 594



(a) (b) (c)

Exam equation sheet, Introduction to statistical inference 2

$\mathsf{E}(\boldsymbol{X}_1\|\boldsymbol{x}_2) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)$	$\mathrm{var}(\boldsymbol{X}_1\|\boldsymbol{x}_2) = \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{21}$
$f(x) = \frac{1}{\beta}e^{-x/\beta}$, $\mathsf{E}(X) = \beta$, $\mathrm{var}(X) = \beta^2$	$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n\|\boldsymbol{\Sigma}\|}}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$
$f_Z(z) = \int_{-\infty}^{\infty} f_X(w)f_Y(z-w)dw$	$(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$
$f(x) = \frac{1}{\Gamma(p/2)2^{p/2}}x^{p/2-1}e^{-x/2}$	$f(t) = \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})}\frac{1}{(p\pi)^{1/2}}\frac{1}{(1+t^2/p)^{(p+1)/2}}$; $\mathsf{E}(T) = 0$; $\mathrm{var}(T) = \frac{p}{p-2}$
$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F(n-1, m-1)$	$f(x) = \frac{\Gamma(\frac{p+q}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{q}{2})}\left(\frac{p}{q}\right)^{p/2}\frac{x^{p/2-1}}{(1+\frac{p}{q}x)^{(p+q)/2}}$
$P(X \geq t) \leq \frac{\mathsf{E}(X)}{t}$	$P(\|X - \mathsf{E}(X)\| \geq t) \leq \frac{\mathrm{var}(X)}{t^2}$
$\lim_{n\to\infty} P(\|X_n - X\| < \epsilon) = 1$	$\lim_{n\to\infty} P(\|X_n - \mu\| < \epsilon) = 1$
If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} a$ then $Y_nX_n \xrightarrow{d} aX$ and $X_n + Y_n \xrightarrow{d} X + a$.	$P(X = x\|r,p) = \binom{r+x-1}{x}p^r(1-p)^x$ $\mathsf{E}(X) = \frac{r(1-p)}{p}$; $\mathrm{var}(X) = \frac{r(1-p)}{p^2}$
$P(X = x\|n,p) = \binom{n}{x}p^x(1-p)^{n-x}$ $\mathsf{E}(X) = p$; $\mathrm{var}(X) = np(1-p)$	$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$
$f(x) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1}e^{-(x/b)^a}$; $E(X^k) = b^k\Gamma(1+k/a)$	$\Gamma(z) = \int_0^{\infty} z^{u-1}e^{-z}du$
$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}$; $\mathsf{E}(X) = \frac{\alpha}{\alpha+\beta}$, $\mathrm{var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$\pi(\theta\|\boldsymbol{x}) = \frac{f(\boldsymbol{x}\|\theta)\pi(\theta)}{m(\boldsymbol{x})}$
$f(x\|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$; $\mathsf{E}(X) = \mathrm{var}(X) = \lambda$	$\mathrm{var}_\theta\, W(\boldsymbol{X}) \geq \frac{\left(\frac{\partial}{\partial\theta}\mathsf{E}_\theta W(\boldsymbol{X})\right)^2}{\mathsf{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\ln f(\boldsymbol{X}\|\theta)\right]^2\right)}$
$\mathsf{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\ln f(\boldsymbol{X}\|\theta)\right]^2\right) = n\mathsf{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\ln f(X\|\theta)\right]^2\right)$	$\mathsf{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\ln f(X\|\theta)\right]^2\right) = -\mathsf{E}_\theta\left(\frac{\partial^2}{\partial\theta^2}\ln f(X\|\theta)\right)$
$\sqrt{n}\left(\tau(\widehat{\theta}) - \tau(\theta)\right) \xrightarrow{d} N(0, v(\theta))$	$\widehat{\mathrm{var}}\tau(\widehat{\theta}) \approx \frac{\left[\tau'(\theta)\right]^2\|_{\theta=\widehat{\theta}}}{-\frac{\partial^2}{\partial\theta^2}\ln L(\theta\|\boldsymbol{x})\|_{\theta=\widehat{\theta}}}$
$[I(\boldsymbol{\theta})]_{i,j} = \mathsf{E}_{\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\theta_i}\ln f(\boldsymbol{X}\|\boldsymbol{\theta})\right)\left(\frac{\partial}{\partial\theta_j}\ln f(\boldsymbol{X}\|\boldsymbol{\theta})\right)\right]$	$e^{-t} = \lim_{n\to\infty}\left(1 - \frac{t}{n}\right)^n$
$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$	$f(x) = 1/(b-a)$; $\mathsf{E}(X) = \frac{b-a}{2}$; $\mathrm{var}\, X = \frac{(b-a)^2}{12}$

```
> qt(0.95,df=13)
[1] 1.770933
> qt(0.975,df=13)
[1] 2.160369
> qt(0.95,df=14)
[1] 1.76131
> qt(0.975,df=14)
[1] 2.144787
> qt(0.95,df=15)
[1] 1.75305
> qt(0.975,df=15)
[1] 2.13145
> qnorm(0.95)
[1] 1.644854
> qnorm(0.975)
[1] 1.959964
> qnorm(0.99)
[1] 2.326348
```