Please provide complete and well-written solutions to the following exercises.

Due February 15, at the beginning of class.

Homework 3

Exercise 1. Let $n \geq 2$ be a positive integer. Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. For any $x, y \in \mathbf{R}^n$, define $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$ and $||x|| := \langle x, x\rangle^{1/2}$. Let $S^{n-1} := \{x \in \mathbf{R}^n \colon ||x|| = 1\}$ be the sphere of radius 1 centered at the origin. Let $x \in S^{n-1}$ be fixed. Let v be a random vector that is uniformly distributed in S^{n-1}.

In modern statistics and data science, data can arise as vectors on high-dimensional spheres. A high-dimensional sphere is rather different from a low-dimensional one, so our intuition about the data in low dimensions may not apply any more in high dimensions. For example, any "equator" of the sphere has most of the mass near it, in the following sense:

For any $t > 0$, and for any $x \in S^{n-1}$ that is fixed,

$$\mathbf{P}(v \in S^{n-1} \colon\ |\langle v, x\rangle| > t/\sqrt{n}) \leq \frac{10}{t}.$$

(Hint: it might be helpful to use Markov's inequality.)

Exercise 2. Let X be uniformly distributed on $[0, 1]$. Show that the location family of X is not an exponential family in the following sense. The corresponding densities $\{f(x+\mu)\}_{\mu\in\mathbf{R}}$ cannot be written in the form

$$h(x)\exp(w(\mu)t(x) - a(w(\mu)))$$

where $h\colon \mathbf{R} \to [0, \infty)$, $w\colon \mathbf{R} \to \mathbf{R}$, $t\colon \mathbf{R} \to \mathbf{R}$, $x \in \mathbf{R}$ and $a(w(\mu))$ is a real number chosen so that the integral of the density is one. (Hint: Argue by contradiction. Assume that the location family is a one-parameter exponential family. Compare where the different densities are zero or nonzero as the parameter changes.)

Exercise 3. Suppose we have a k-parameter exponential family in canonical form so that

$$f_w(x) := h(x)\exp\Big(\sum_{i=1}^k w_i t_i(x) - a(w)\Big),$$

$$a(w) := \log \int_{\mathbf{R}^n} h(x)\exp\Big(\sum_{i=1}^k w_i t_i(x)\Big) d\mu(x).$$

$$W := \{w \in \mathbf{R}^k \colon a(w) < \infty\}.$$

Show that $a(w)$ is a convex function. That is, for any $w_1, w_2 \in \mathbf{R}^k$ and for any $t \in (0, 1)$,

$$a(tw_1 + (1-t)w_2) \leq ta(w_1) + (1-t)a(w_2).$$

(Hint: use Hölder's inequality of the form $\int |fg|\, d\mu \leq (\int |f|^p\, d\mu)^{1/p}(\int |g|^q\, d\mu)^{1/q}$ where $1/p + 1/q = 1$, where $p = t^{-1}$.)

Conclude that the set W is a convex set. (That is, if $w_1, w_2 \in W$ then for any $t \in [0,1]$, $tw_1 + (1-t)w_2 \in W$.)

Exercise 4. Using a two parameter exponential family for a Gaussian random variable (with mean μ and variance σ^2), compute both sides of the following identity in terms of μ and σ:

$$e^{-a(w)} \frac{\partial^2}{\partial w_i \partial w_j} e^{a(w)} = \int_{\mathbf{R}} t_i(x) t_j(x) h(x) \exp\Big(\sum_{i=1}^{2} w_i t_i(x) - a(w)\Big) d\mu(x), \qquad \forall\, 1 \leq i,j \leq 2.$$

Recall that in this case,

$$t_1(x) := x, \qquad t_2(x) := x^2, \qquad w_1 := \frac{\mu}{\sigma^2}, \qquad w_2 := -\frac{1}{2\sigma^2},$$
$$a(w) := -\frac{w_1^2}{4w_2} - \frac{1}{2}\log(-2w_2).$$

Exercise 5. Let $X\colon \Omega \to \mathbf{R}^n$ be a random variable with the **standard Gaussian distribution**:

$$\mathbf{P}(X \in A) := \int_A e^{-(x_1^2 + \cdots + x_n^2)/2} dx (2\pi)^{-n/2}, \qquad \forall\, A \subseteq \mathbf{R}^n \text{ measurable.}$$

Let $v_1, \ldots, v_m$ be vectors in $\mathbf{R}^n$. Let $\langle \cdot, \cdot \rangle : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ be the standard inner product on $\mathbf{R}^n$, so that $\langle x, y \rangle := \sum_{i=1}^n x_i y_i$ for any $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbf{R}^n$.

First, let $v \in \mathbf{R}^n$ and show that $\langle X, v \rangle$ is a mean zero Gaussian with variance $\langle v, v \rangle$.

Then, show that the random variables $\langle X, v_1 \rangle, \ldots, \langle X, v_m \rangle$ are independent if and only if the vectors $v_1, \ldots, v_m$ are pairwise orthogonal.

(Hint: use the rotation invariance of the Gaussian.)

Exercise 6. Recall that the gamma distribution has two parameters $\alpha, \beta > 0$.

- Show that the gamma distribution is a 2-parameter exponential family.
- Find the mean and variance of a gamma distributed random variable by differentiating the exponential family.
- Find the moment generating function of a gamma distributed random variable, and use it to find the distribution of $\sum_{i=1}^n X_i$ where $X_1, \ldots, X_n$ are independent, and X_i has gamma distribution with parameters α_i and β for all $1 \leq i \leq n$.

  You may use without proof the following uniqueness result about moment generating functions (MGFs): If Y and Z are two random variables whose MGFs coincide in a neighborhood of 0 ($\exists \delta > 0$ for which $M_Y(u) = M_Z(u) < \infty$ for all $u \in [-\delta, \delta]$), then Y and Z have the same distribution.

Exercise 7. Let $n \geq 2$ be an integer. Let $X_1, \ldots, X_n$ be a random sample of size n. Assume that $\mu := \mathbf{E}X \in \mathbf{R}$ and $\sigma := \sqrt{\mathrm{var}(X)} < \infty$. Let $\overline{X}$ be the sample mean and let S be the sample standard deviation of the random sample. Show the following

- $\text{Var}(\overline{X}) = \sigma^2/n$.
- $\mathbf{E}S^2 = \sigma^2$.

Exercise 8. Let $X\colon \Omega \to \mathbf{R}$ be a random variable with $\mathbf{E}X^2 < \infty$. Show that the quantity $\mathbf{E}(X-t)^2$ is minimized for $t \in \mathbf{R}$ uniquely when $t = \mathbf{E}X$.

Exercise 9. Let X be a chi squared random variables with p degrees of freedom. Let Y be a chi squared random variable with q degrees of freedom. Assume that X and Y are independent. Show that $(X/p)/(Y/q)$ has the following density, known as **Snedecor's f-distribution** with p and q degrees of freedom

$$f_{(X/p)/(Y/q)}(t) := \frac{t^{(p/2)-1}(p/q)^{p/2}\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)}\Big(1 + t(p/q)\Big)^{-(p+q)/2}, \qquad \forall\, t > 0.$$