Homework #1

Due: 2025-10-12 23:59 | 8 Problems, 100 Pts

Name: 赵凌哲, ID: 2300015881

Problem 1 (8'). Define

$$\Gamma(s) = \int_0^{+\infty} x^{s-1}\mathrm{e}^{-x}\mathrm{d}x.$$

where $s > 0$. It can be proved that $\Gamma(s)$ is well-defined (You don't need to prove this).

(1) (4') Prove that, $\Gamma(s+1) = s\Gamma(s)$.

(2) (4') Prove that,

$$\Gamma(s) = 2\int_0^{+\infty} x^{2s-1}\mathrm{e}^{-x^2}\mathrm{d}x.$$

◀

Answer.

(1)

$$\begin{aligned}
\Gamma(s+1) &= \int_0^{+\infty} x^s\mathrm{e}^{-x}\mathrm{d}x \\
&= -\int_0^{+\infty} x^s\mathrm{d}\mathrm{e}^{-x} \\
&= -x^s\mathrm{e}^{-x}\Big|_0^{+\infty} + \int_0^{+\infty} \mathrm{e}^{-x}\mathrm{d}x^s \\
&= 0 + s\int_0^{+\infty} x^{s-1}\mathrm{e}^{-x}\mathrm{d}x \\
&= s\Gamma(s).
\end{aligned}$$

(2)

$$\begin{aligned}
\Gamma(s) &= \int_0^{+\infty} x^{s-1}\mathrm{e}^{-x}\mathrm{d}x \\
&\overset{x=t^2}{=} \int_0^{+\infty} (t^2)^{s-1}\mathrm{e}^{-t^2}\mathrm{d}t^2 \\
&= \int_0^{+\infty} 2t\cdot t^{2s-2}\mathrm{e}^{-t^2}\mathrm{d}t \\
&= 2\int_0^{+\infty} x^{2s-1}\mathrm{e}^{-x^2}\mathrm{d}x.
\end{aligned}$$

◁

Problem 2 (10'). Define a random variable $Q \sim \chi^2(k)$ $(k \in \mathbb{Z}_+)$ if $Q = Z_1^2 + \cdots + Z_k^2$ where $Z_1, \cdots, Z_k \sim \mathcal{N}(0,1)$ are independent random variables. Given two independent random variables $X \sim \chi^2(m),\ Y \sim \chi^2(n)$ $(m, n \in \mathbb{Z}_+,\ m > n)$.

(1) (4') Calculate the value of $\mathbb{E}(X)$.

(2) (2') Prove that, $X+Y \sim \chi^2(m+n)$.

(3) (4') Does $X-Y \sim \chi^2(m-n)$? Prove your result.

◀

Answer.

(1) 因为 $X \sim \chi^2(m)$，所以可设 $X = Z_1^2 + \cdots + Z_m^2$ where $Z_1, \cdots, Z_m \sim \mathcal{N}(0,1)$ are independent random variables.

$$
\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(Z_1^2 + \cdots + Z_m^2) \\
&= \sum_{i=1}^{m} \mathbb{E}(Z_i^2) \\
&= \sum_{i=1}^{m} (Var(Z_i) + [\mathbb{E}(Z_i)]^2) \\
&= \sum_{i=1}^{m} 1 \\
&= m.
\end{aligned}
$$

(2) 因为 $X \sim \chi^2(m)$，所以可设 $X = U_1^2 + \cdots + U_m^2$ where $U_1, \cdots, U_m \sim \mathcal{N}(0,1)$ are independent random variables.
又因为 $Y \sim \chi^2(n)$，所以可设 $Y = V_1^2 + \cdots + V_n^2$ where $V_1, \cdots, V_n \sim \mathcal{N}(0,1)$ are independent random variables.
又因为 X, Y 相互独立，从而 $\{U_1, \cdots, U_m\}, \{V_1, \cdots, V_n\}$ 是相互独立的两组标准正态随机变量，因此：

$$
X + Y = (U_1^2 + \cdots + U_m^2) + (V_1^2 + \cdots + V_n^2).
$$

根据卡方分布的定义，我们有：

$$
X + Y \sim \chi^2(m+n).
$$

(3) 否。我们设：

$$
Z \sim \chi^2(m-n).
$$

如果有 $X - Y \sim \chi^2(m-n)$，则有：

$$
Var(X-Y) = Var(Z).
$$

下面证明其错误：

$$\begin{aligned} Var(Z) &= Var(W_1^2 + \cdots + W_{m-n}^2) \\ &= \sum_{i=1}^{m-n} Var(W_i^2) \qquad (\textit{All of } W_i \textit{ are independent}) \\ &= (m-n)Var(W^2), \quad \textit{where } W \sim \mathcal{N}(0,1). \end{aligned}$$

$$\begin{aligned} Var(X-Y) &= Var((U_1^2 + \cdots + U_m^2) - (V_1^2 + \cdots + V_n^2)) \\ &= \sum_{i=1}^{m} Var(U_i^2) + \sum_{i=1}^{n} Var(V_i^2) \qquad (X, Y \textit{ are independent}) \\ &= mVar(U^2) + nVar(V^2), \quad \textit{where } U, V \sim \mathcal{N}(0,1). \end{aligned}$$

因为 $U, V, W \sim \mathcal{N}(0,1).$，从而有 $Var(U^2) = Var(V^2) = Var(W^2) = k \neq 0$（实际上可计算 $k = 2$），因此：

$$Var(Z) = (m-n)k \neq (m+n)k = Var(X-Y).$$

从而有 $X - Y$ 不服从 $\chi^2(m-n)$.（其实直接用 $X - Y$ 可能为负也可证明）

◁

Problem 3 (8'). In this problem, we will prove two basic probability inequalities. You will use these inequalities frequently in this course, so make sure you are familiar with them, including the statements and conditions.

(1) (4') (Markov Inequality) Suppose X is a non-negative random variable and $\mathbb{E}(X) < +\infty$. Prove that, for any $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

(2) (4') (Chebyshev Inequality) Suppose X is a random variable with $\mathbb{E}(X) < +\infty, Var(X) < +\infty$. Prove that, for any $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{Var(X)}{a^2}.$$

◀

Answer.

(1) 因为 $a \geq 0$，所以有：

$$\begin{aligned} \mathbb{E}(X) &\geq \mathbb{E}(X \cdot I_{X \geq a}) \\ &\geq \mathbb{E}(a \cdot I_{X \geq a}) \\ &= a\mathbb{E}(I_{X \geq a}) \\ &= a\mathbb{P}(X \geq a). \end{aligned}$$

从而 Markov Inequality 得证.

(2) 令 $Y=(X-\mathbb{E}(X))^2$，因为 $\mathbb{E}(X)<+\infty, Var(X)<+\infty$，从而 $\mathbb{E}(Y)<+\infty$，对其使用 Markov Inequality：

$$\begin{aligned}\mathbb{P}(Y\geq a^2)&=\mathbb{P}(|X-\mathbb{E}(X)|\geq a)\\&\leq\frac{\mathbb{E}(Y)}{a^2}\\&=\frac{Var(X)}{a^2}.\end{aligned}$$

从而 Chebyshev Inequality 得证.

◁

Problem 4 (14'). In this problem, we will prove the following Chernoff bound (Theorem 1).

Theorem 1 (Chernoff Bound)**.** *Suppose $X_1,\cdots,X_n$ are independently Bernoulli random variables with expectation $p\in(0,1)$. Then for any $\varepsilon\in(0,1-p)$,*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i\geq p+\varepsilon\right)\leq\exp\left[-nD_B(p+\varepsilon||p)\right]\leq\exp(-2n\varepsilon^2),$$

where

$$D_B(p||q):=p\ln\frac{p}{q}+(1-p)\ln\frac{1-p}{1-q}.$$

(1) (5') Under the conditions of Theorem 1, prove that for any $t>0$,

$$\mathbb{P}\left(\sum_{i=1}^n X_i\geq n(p+\varepsilon)\right)\leq\mathbb{E}\left(\mathrm{e}^{t\sum_{i=1}^n X_i}\right)\cdot\mathrm{e}^{-nt(p+\varepsilon)}.$$

(2) (9') Finish the proof of Theorem 1.

◀

Answer.

(1) 对于 any $t>0$，我们由 Markov Inequality 有：

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^n X_i\geq n(p+\varepsilon)\right)&=\mathbb{P}\left(\mathrm{e}^{t\sum_{i=1}^n X_i}\geq\mathrm{e}^{nt(p+\varepsilon)}\right)\\&\leq\mathbb{E}\left(\mathrm{e}^{t\sum_{i=1}^n X_i}\right)\cdot\mathrm{e}^{-nt(p+\varepsilon)}.\end{aligned}$$

从而所给不等式得证.

(2) 因为上一问的结论对于 any $t > 0$ 成立，从而有：

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq p+\varepsilon\right) \leq \inf_{t>0}\ \mathbb{E}\left(\mathrm{e}^{t\sum_{i=1}^{n} X_i}\right)\cdot \mathrm{e}^{-nt(p+\varepsilon)}$$

其中：

$$\mathbb{E}\left(\mathrm{e}^{t\sum_{i=1}^{n} X_i}\right) = \prod_{i=1}^{n}\mathbb{E}[e^{tX_i}] = \left(\mathbb{E}[e^{tX}]\right)^n = \left(pe^t + (1-p)\right)^n$$

因此：

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq p+\varepsilon\right) \leq \inf_{t>0}\ \left(e^{-nt(p+\varepsilon)}\cdot \left(pe^t + (1-p)\right)^n\right)$$

不妨令

$$g(t) = \frac{1}{n}\ln[e^{-n(p+\varepsilon)t}\cdot(pe^t+(1-p))^n] = -(p+\varepsilon)t + \ln\left(pe^t + (1-p)\right)$$

求导得

$$g'(t) = \frac{d}{dt}\left[-(p+\varepsilon)t + \ln\left(pe^t+(1-p)\right)\right] = -(p+\varepsilon) + \frac{pe^t}{pe^t+(1-p)}$$

令 $g'(t^*) = 0$，求解得最优参数 t^*:

$$-(p+\varepsilon) + \frac{pe^{t^*}}{pe^{t^*}+(1-p)} = 0$$

$$e^{t^*} = \frac{(p+\varepsilon)(1-p)}{p(1-p-\varepsilon)} \quad (\text{题设}\varepsilon < 1-p)$$

代入原式 $g(t)$，

$$g(t^*) = -(p+\varepsilon)\ln\left(\frac{p+\varepsilon}{p}\right) + (1-p-\varepsilon)\ln\left(\frac{1-p}{1-p-\varepsilon}\right)$$

又因为

$$(1-p-\varepsilon)\ln\left(\frac{1-p}{1-p-\varepsilon}\right) = -(1-p-\varepsilon)\ln\left(\frac{1-p-\varepsilon}{1-p}\right)$$

则

$$g(t^*) = -\left[(p+\varepsilon)\ln\left(\frac{p+\varepsilon}{p}\right) + (1-p-\varepsilon)\ln\left(\frac{1-p-\varepsilon}{1-p}\right)\right] = -D_{\mathrm{B}}(p+\varepsilon \parallel p)$$

由二阶近似，可求得右式最大值：

$$\begin{aligned} -D_B(p+\varepsilon \parallel p) &\xrightarrow{\epsilon\to 0} -\frac{\varepsilon^2}{2p(1-p)} \\ &\leq -2\varepsilon^2. \qquad (when\ p = \frac{1}{2}) \end{aligned}$$

从而有：

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq p+\varepsilon\right) \leq \exp\left[-nD_B(p+\varepsilon||p)\right] \leq \exp(-2n\varepsilon^2).$$

◁

Problem 5 (12').

(1) (3') For what value of d does the volume $V(d)$ of a d-dimensional unit ball take on its maximum? You don't need to prove your result.

(2) (4') Consider drawing a random point x from the unit ball in $\mathbb{R}^d$ (surface and interior) uniformly at random. What's the variance of x_1 (the first coordinate of x)? You don't need to prove your result.

(3) (5') Recall the way we generate a random unit vector in d dimensions. First, generate d i.i.d samples v_i ($i \in [d]$) from a Gaussian distribution with $\mu = 0$ and $\sigma^2 = 1$. Then, define the vector v as $v = [v_1, v_2, ..., v_d]$. Finally, return $\frac{v}{||v||_2}$, where $||v||_2 = \sqrt{\sum_{i=1}^d v_i^2}$. Briefly explain why it generates a unit vector uniformly at random.

◀

Answer.

(1)

$$V(d) = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}.$$

当 $d = 5$ 时取最大.

(2)

$$Var(x_1) = \frac{1}{d} \cdot \frac{d}{d+2} = \frac{1}{d+2}.$$

(3) 对于 $v \sim \mathcal{N}(0, I_d)$，其密度函数只依赖于 $||v||_2$，即具有旋转不变性.
因此对于 $\frac{v}{||v||_2}$，我们就得到了具有旋转不变性的单位向量（均匀测度），即为 generate a unit vector uniformly at random.

◁

Problem 6 (12'). A 3-dimensional cube has vertices, edges and faces. In a d-dimensional cube, these components are called faces. A vertex is a 0-dimensional face, an edge a 1-dimensional face, etc. Answer the following problems. You don't need to prove your result.

(1) (3') For $0 \le k \le d$, how many k-dimensional faces does a d-dimensional cube have?

(2) (3') What is the total number of faces of all dimensions? The d-dimensional face is the cube itself which you can include in your count.

(3) (3') What is the surface area of a unit cube in d-dimensions (a unit cube has side-length one in each dimension)?

(4) (3') What is the surface area of the cube if the length of each side was 2?

◀

Answer.

(1)

$$\binom{d}{k} \cdot 2^{d-k}.$$

(2)

$$\sum_{k=0}^{d} \binom{d}{k} \cdot 2^{d-k} = 3^d.$$

(3)

$$2d.$$

(4)

$$2d \cdot 2^{d-1} = d \cdot 2^d.$$

◁

Problem 7 (16'). Consider the upper hemisphere of a unit-radius ball in d-dimensions. What is the height of the maximum volume cylinder that can be placed entirely inside the hemisphere?

[Hint: You need to consider all possible placement options.] ◀

Answer. 考虑如下若干放置方法：

- 圆柱 (d-1) 维的底面置于 d 维半球的底面.
- 圆柱 1 维的母线（长度为高）置于 d 维半球的底面.
- 其余倾斜（不对称）的放置方法.

对于第一种情况，有：

$$V = h \cdot (1-h^2)^{\frac{d-1}{2}} \cdot V(d-1).$$

求导可以得到使其最大的 h 和此时的 V：

$$h = \sqrt{\frac{1}{d}},$$
$$V = \sqrt{\frac{1}{d}} \cdot \left(\sqrt{\frac{d-1}{d}}\right)^{d-1} \cdot V(d-1).$$

对于第二种情况，有：

$$V = h \cdot \left(\frac{\sqrt{1-\frac{h^2}{4}}}{2}\right)^{d-1} \cdot V(d-1).$$

求导可以得到使其最大的 h 和此时的 V：

$$h = \sqrt{\frac{2}{d}},$$
$$V = \frac{1}{2^{d-2}} \cdot \sqrt{\frac{1}{d}} \cdot \left(\sqrt{\frac{d-1}{d}}\right)^{d-1} \cdot V(d-1).$$

对于第三种情况，因为其不对称，直观来看，因为整体约束是对称的，极值解应当具有对称形式. 否则会出现对称的无数极值点。或者说可以将不对称情况通过微小调整变为对称，且保证体积不减小.

综上讨论，认为第一种情况可以取圆柱体积最大值，此时其高为 $\sqrt{\frac{1}{d}}$.

◁

Problem 8 (20'). Consider the following geometries in d-dimensional space:

$$\Phi_d = \left\{(x_1, \cdots, x_d) \mid x_1^2 + \cdots + x_d^2 \le 1\right\} \qquad \Omega_d = \{(x_1, \cdots, x_d) \mid |x_1| + \cdots + |x_d| \le 1\}$$

(1) (6') Consider the following two random processes:

- Pick two uniformly random unit vectors x, y from the sphere in d dimension.
- Pick a uniformly random plane passing through the origin in d dimensions. Then pick two uniformly random unit vectors w, z from the 2D sphere lying on that plane.

Determine whether x and w are identically distributed, and whether the pairs (x, y) and (w, z) are identically distributed. Briefly explain the reasons.

(2) (6') Suppose x and y are sampled independently from Φ_d. Denote random variable $Z = x^\top y$. Calculate $Var(Z)$.

(3) (8') Calculate $V(\Omega_d)$ (the volume of Ω_d) and $A(\Omega_d)$ (the surface area of Ω_d).

◀

Answer.

(1) x and w are identically distributed. 首先我们知道 x 具有旋转不变性（均匀分布），对于 w，我们记任意旋转为 Q，取平面的操作为 P，则有：

$$P \overset{d}{=} QP,$$
$$w|P \overset{d}{=} Qw|QP.$$

从而 $w \stackrel{d}{=} Qw$，w 也具有旋转不变性，因此 x and w are identically distributed.

The pairs (x, y) and (w, z) are not identically distributed. 我们只需要研究 pair 中两个向量的夹角：

- x, y 的夹角并非均匀分布，固定 x 为极点，y 靠近赤道位置的概率大.（表达式见下一问）
- w, z 取自同一个二维平面，因此夹角满足均匀分布.

因此仅从夹角的概率分布即可看出 pairs (x, y) and (w, z) are not identically distributed.

(2) 我们知道：

$$Z = ||x||_2 \cdot ||y||_2 \cdot cos\theta.$$

因此：

$$\begin{aligned} E(Z) &= 0, \\ Var(Z) &= E(Z^2) \\ &= E\left[(||x||_2 \cdot ||y||_2 \cdot cos\theta)^2\right] \\ &= E(r^2)^2 \cdot E(cos^2\theta). \end{aligned}$$

根据 Problem 5 中第 (2) 问得出答案的过程，我们有：

$$\begin{aligned} E(r^2) &= d \cdot E(x_1^2) \\ &= \frac{d}{d+2}. \end{aligned}$$

而如果设 $I_k = \int_0^\pi (\sin x)^k \, dx$，则有：

$$\begin{aligned} f(\theta) &\propto (\sin\theta)^{d-2}, \\ f(\theta) &= \frac{1}{I_{k-2}} \cdot (\sin\theta)^{d-2}. \end{aligned}$$

从而有：

$$\begin{aligned} E(cos^2\theta) &= 1 - E(sin^2\theta) \\ &= 1 - \frac{1}{I_{d-2}} \cdot \int_0^\pi (\sin x)^2 \cdot (\sin x)^{d-2} \, dx \\ &= 1 - \frac{I_d}{I_{d-2}} \\ &= 1 - \frac{d-1}{d} \\ &= \frac{1}{d}. \end{aligned}$$

于是有:

$$\begin{aligned}
Var(Z) &= E(r^2)^2 \cdot E(cos^2\theta) \\
&= \left(\frac{d}{d+2}\right)^2 \cdot \frac{1}{d} \\
&= \frac{d}{(d+2)^2}.
\end{aligned}$$

(3) Ω_d 可以按照各个 x_i 的正负分为完全相同的 2^d 个部分. 只需考虑全部 $x_i \geq 0$ 的部分. 下面用数学归纳法证明其体积为 $\frac{1}{d!}$:
当 $d = 2, 3$ 时，容易求出 V_d 分别为 $\frac{1}{2}, \frac{1}{6}$.
假设 $d = k-1$ 时，$V_{k-1} = \frac{1}{(k-1)!}$，则有:

$$\begin{aligned}
V_k &= \int_0^1 x^{k-1} \cdot V_{k-1}\, dx \\
&= \frac{1}{k} \cdot V_{k-1} \\
&= \frac{1}{k!}.
\end{aligned}$$

从而证出每部分的体积为 $\frac{1}{d!}$. 总体积为:

$$V(\Omega_d) = \frac{2^d}{d!}.$$

对于上面求出的每个部分的体积 V_k，注意到还有另外一种求法:

$$\begin{aligned}
V_k &= \int_0^{\frac{1}{\sqrt{k}}} \frac{x^{k-1}}{\left(\frac{1}{\sqrt{k}}\right)^{k-1}} \cdot A_k\, dx \\
&= \frac{1}{k} \cdot \frac{\left(\frac{1}{\sqrt{k}}\right)^{k}}{\left(\frac{1}{\sqrt{k}}\right)^{k-1}} \cdot A_k \\
&= \frac{1}{k \cdot \sqrt{k}} \cdot A_k \\
&= \frac{1}{k!}.
\end{aligned}$$

这种方法是按照垂直于该部分表面的轴进行积分，每个切面都是放缩的 A_k.
从而可以解得:

$$A_k = \frac{\sqrt{k}}{(k-1)!}.$$

于是有:

$$A(\Omega_d) = \frac{2^d \cdot \sqrt{d}}{(d-1)!}.$$

◁