1. Total: 85/100 (A-).

(a) Prove: $\inf_{\lambda>0} M_X(\lambda) e^{-\lambda u} \geq \inf_{k \in \mathbb{Z}^+} [\mathbb{E}X^k] u^{-k}$

Proof:

First, let me show that $\mathbb{E}X^k \leq \left(\frac{k}{\lambda e}\right)^k M_X(\lambda)$

I found this result on wiki that's OK if you understand it (point for honesty!)

Proving the claim ...

given $X > 0$ a.s, $k \geq 0$, $\lambda > 0$

using $1+x \leq e^x \Rightarrow 1 + \frac{\lambda x}{k} \leq e^{\lambda x/k - 1}$ cannot be there both at the same time

Take $t = \frac{\lambda x}{k}$, then $t = 1 + (t-1) \leq e^{t-1}$

$\Leftrightarrow \frac{\lambda x}{k} \leq e^{\frac{\lambda x}{k} - 1}$.

$x^k \leq \left(\frac{k}{\lambda e}\right)^k e^{\lambda x}$.

$\Rightarrow x^k \leq \left(\frac{k}{\lambda e}\right)^k e^{\lambda x}$.

wrong implication (even if the premise held)

Now, ... rearranging we have

$$x^k \leq \left(\frac{k}{\lambda e}\right)^k e^{\lambda x}.$$

[mistakes — not sure you got the proof.]

not sure what you do...

Golden rule (by A. Juditsky): "Never use sth you haven't proved."

$$\Rightarrow \quad \mathbb{E} X^k \le \left(\frac{k}{\lambda e}\right)^k \mathbb{E} e^{\lambda X}$$

+/−

$$\Rightarrow \inf_{k \in \mathbb{Z}^+} \mathbb{E} X^k \le \left(\frac{k}{\lambda e}\right)^k \mathbb{E} e^{\lambda X}$$

Multiply by u^{-k} both sides, $\because u, k \ge 0$

$$\Rightarrow \inf_{k \in \mathbb{Z}^+} \mathbb{E} X^k u^{-k} \le \left(\frac{k}{\lambda e u}\right)^k M_X(\lambda).$$

Choose $k = \lambda u$ on the RHS

$$\Rightarrow \text{RHS} = \left(\frac{\lambda u}{\lambda e u}\right)^{\lambda u} M_X(\lambda)$$

$$= e^{-\lambda u} M_X(\lambda).$$

Since $\inf_{k \in \mathbb{Z}^+} \mathbb{E} X^k \le e^{-\lambda u} M_X(\lambda)$ holds $\forall \lambda > 0$.

we can take $\inf \lambda$ on R.H.S

☑

OK, this inequality indeed implies (a)

(b) Given that X is symmetric.

We can start off with $X = X^+ + X^-$.

where $X^+ := \max\{X, 0\}$, $X^- = \max\{-X, 0\}$.

$$\underbrace{X^+}_{\geq 0 \text{ a.s}} \qquad \underbrace{X^-}_{\leq 0 \text{ a.s}}$$

Now:

$$M_X(\lambda) e^{-\lambda u} = \mathbb{E} e^{\lambda X} \cdot e^{-\lambda u} = \mathbb{E} e^{\lambda(X^+ + X^-)} \cdot e^{-\lambda u}$$

$$= \mathbb{E}\left[e^{\lambda X^+} \cdot e^{\lambda X^-}\right] e^{-\lambda u} \underbrace{\geq}_{\text{AM-GM ineq.}} \mathbb{E}\left[\frac{e^{\lambda X^+} + e^{\lambda X^-}}{2}\right] e^{-\lambda u}$$

$$= \frac{1}{2}\left(\mathbb{E} e^{\lambda X^+} \cdot e^{-\lambda u} + \mathbb{E} e^{\lambda X^-} \cdot e^{-\lambda u}\right)$$

$$= \frac{1}{2} \mathbb{E}\left(e^{\lambda X^+} + e^{\lambda X^-}\right) \cdot e^{-\lambda u}$$

By symmetry of X, $\mathbb{E} X^+ = -\mathbb{E} X^-$

or more generally, $\mathbb{E}(X^+)^m = -\mathbb{E}(X^-)^m$

for m being odd number.

Correct :)

$$\therefore \mathbb{E}\frac{e^{\lambda\frac{x^+}{u}} + e^{\lambda\frac{x^-}{u}}}{2} = \cancel{\frac{1}{2}\mathbb{E}e^{\lambda\frac{x^2}{u^2}}}, \quad u > 0$$

The above can be proven by expanding the Taylor Series & cancelling the odd terms: If you actually did this, you would see that the even moments don't match.

We now have $M_x(\lambda)e^{-\lambda u} \geq \frac{1}{2}\mathbb{E}e^{\lambda x^2}\cdot e^{-\lambda u}$

$$\Rightarrow \inf_{\lambda>0} M_x(\lambda)e^{-\lambda u} \geq \inf(\lambda>0)\,\frac{1}{2}\mathbb{E}e^{\lambda x^2}\cdot e^{-\lambda u}.$$

$$= \frac{1}{2}\inf M_{x^2}(\lambda)\,e^{-\lambda u}.$$

Btw, note that

$$\frac{e^{X_+}+e^{X_-}}{2} = \frac{e^{X}+e^{-X}}{2}$$

$$\geq \frac{1}{2}\inf_{k\in\mathbb{Z}^+}\mathbb{E}[X^{2k}]\,u^{-2k}$$

(from (a))

. $\square$

+/–

2.) To prove $K_X(\lambda t_1 + (1-\lambda) t_2) \le \lambda K_X(t_1) + (1-\lambda) K_X(t_2)$

Proof;

$$\text{LHS} = \log \mathbb{E}\left[e^{(\lambda t_1 + (1-\lambda) t_2) x}\right] = \log \mathbb{E}\left[e^{\lambda t_1 x} \cdot e^{(1-\lambda) t_2 x}\right]$$

$$= \log \sum_{i \in [n]} \left(p_i e^{\lambda t_1 x_i} \cdot e^{(1-\lambda) t_2 x_i}\right)$$

$$= \log \sum_{i \in \Omega} \left(p_i^{\lambda} e^{\lambda t_1 x_i} \cdot p_i^{1-\lambda} e^{(1-\lambda) t_2 x_i}\right)$$

$$= \log \sum_{i \in \Omega} \left(\left(p_i e^{t_1 x_i}\right)^{\lambda} \left(p_i e^{t_2 x_i}\right)^{1-\lambda} \right)$$

$$\cancel{\le \log \left[\left(\sum_i \left(p_i e^{t_1 x_i}\right)^{\lambda}\right)^{\lambda/2} \sum_j \left(p_j e^{t_2 x_j}\right)^{\frac{1-\lambda}{2}} \right]} \quad \cancel{[\text{Hint}]}$$

$$\le \lambda \log \sum_i p_i e^{t_1 x_i} + (1-\lambda) \log \sum_j p_j e^{t_2 x_j}$$

$$= \lambda \log \mathbb{E} e^{t_1 X} + (1-\lambda) \log \mathbb{E} e^{t_2 X}. \quad \square$$

$$\sum_i \alpha_i^{\lambda} \beta_i^{1-\lambda} \le \left(\sum_i \alpha_i\right)^{\lambda} \left(\sum_i \beta_i\right)^{1-\lambda} \qquad +$$

Hölder with $p = 1/\lambda$, $q = \frac{1}{1-\lambda}$.

3.)

3.1)

a) $\bar{\Phi}(t) = \int_t^\infty \phi(t)\,dt = \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt$

$$= \int_{t=u}^\infty \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{t} \cdot t e^{-t^2/2}\,dt = -\int_{t=u}^\infty \frac{1}{t}\phi'(t)\,dt$$

$$= -\left[\frac{1}{t}\phi(t)\Big|_{t=u}^\infty + \int_{t=u}^\infty \frac{\phi(t)}{t^2}\,dt\right]$$

$$= \frac{\phi(u)}{u} - \underbrace{\int_{t=u}^\infty \frac{\phi(t)}{t^2}\,dt}_{>0} \leq \frac{\Phi(u)}{u}.$$

+

Now continuing to integrate by parts

$$\int_{t=u}^{\infty} \frac{\phi(t)}{t^2}\,dt = -\int_{t=u}^{\infty} \frac{1}{t^3}\,\phi''(t)\,dt$$

Now

$$\Phi(u) = \frac{\phi(u)}{u} + \int_{t=u}^{\infty} \frac{\phi'(t)}{t^3}\,dt$$

$$= \frac{\phi(u)}{u} - \frac{\phi(u)}{u^3} + 3\underbrace{\int_{t=u}^{\infty} \frac{\phi(t)}{t^4}\,dt}_{\geq 0}$$

$$\geq \frac{\phi(u)}{u} - \frac{\phi(u)}{u^3}$$

$\dagger$

$\blacksquare$

(b) keep continuing: (& continue keeping
— I like this ;)

$$\Phi(u) = \phi(u)\left(\frac{1}{u} - \frac{1}{u^3}\right) - 3\left\{\left[\frac{\phi(t)}{t^5}\right]_u^\infty + 5\int_u^\infty \frac{\phi(t)}{t^6}\,dt\right\}$$

$$= \phi(u)\left(\frac{1}{u} - \frac{1}{u^3} + \frac{3}{u^5}\right) - 15\int_{t=u}^\infty \frac{\phi(t)}{t^6}\,dt$$

$$\leq \phi(u)\left(\frac{1}{u} - \frac{1}{u^3} + \frac{3}{u^5}\right)$$

+

☐

3.2) Power series can be shown

− by induction on k

4

(i) Using the same trick as we did with proving Markov's inequality.

$$\mathbb{E}X = \mathbb{E}\left[X \cdot \mathbb{1}\{X \le (1-t)\,\mathbb{E}X\}\right] + \mathbb{E}\left[X \cdot \mathbb{1}\{X > (1-t)\,\mathbb{E}X\}\right]$$

$\downarrow$ Cauchy-Schwatz:

$$\le (1-t)\,\mathbb{E}X + (\mathbb{E}X^2)^{1/2}\left(\mathbb{P}\{X > (1-t)\,\mathbb{E}X\}\right)^{1/2}.$$

Rearranging:

$$\Rightarrow \mathbb{P}\{X \geq (1-t)\,\mathbb{E}X\} \geq \frac{t[\mathbb{E}X]^2}{\mathbb{E}X^2}$$

+

$\square$.

(iii) To show:

$$\mathbb{P}(X \geq (1-t)\mathbb{E}X) \geq \left(\frac{t^p (\mathbb{E}X)^p}{\mathbb{E}X^p}\right)^{1/p-1}$$

Proof:

Same procedure as (i), but instead of Cauchy Schwartz. use holder:

$$\mathbb{E}X = \mathbb{E}\left[X \cdot \mathbb{1}(X \leq (1-t)\mathbb{E}X)\right]$$

$$+ \mathbb{E}\left[X \cdot \mathbb{1}(X \geq (1-t)\mathbb{E}X)\right] \quad \text{Holder's}$$

$$\leq (1-t)\mathbb{E}X + (\mathbb{E}X^p)^{1/p}\{\mathbb{P}(X \geq (1-t)\mathbb{E}X\}^{\cancel{1/p}}$$

(should be $\frac{1}{q}$).

$$\Rightarrow \quad \mathbb{P}(X \geq (1-t)\mathbb{E}X) \geq \left(\frac{t^p (\mathbb{E}X)^p}{\mathbb{E}X^p}\right)^{\frac{1}{p-1}}$$

+0

$\blacksquare$

5:>

(a) $M_2(t) = \mathbb{E}\left[e^{t(Z_1^2+Z_2^2)}\right]$

$= \mathbb{E}\left[e^{tZ_1^2}\right] \mathbb{E}\left[e^{tZ_2^2}\right]$ (iid)

$\mathbb{E}\, e^{tZ_1^2} = \int e^{tZ_1^2} f_{Z_1}(z_1)\, dz_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz_1^2} e^{-z_1^2/2} dz_1$

$= \frac{1}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} e^{-(\frac{1}{2}-t)z_1^2} dz_1}_{\text{evaluated using well known form for } \int_{-\infty}^{\infty} e^{-kx^2}dx,\ k>0}$

$= \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{2\pi}}{\sqrt{1-2t}}$

$= \frac{1}{\sqrt{1-2t}}, \quad t < 1/2$

+

$\Rightarrow \mathbb{E}\, e^{tZ_1^2}\, \mathbb{E}\, e^{tZ_2^2} = \frac{1}{1-2t}, \quad t < 1/2$

$\Rightarrow M_{2d}(t) = \prod_{i=1}^{d} \mathbb{E}\, e^{tZ_i^2} = \frac{1}{(1-2t)^d}$

+

b) $\mathbb{P}(X > x) = \inf_{t<1/2} \dfrac{e^{-tx}}{(1-2t)^d}$, $x > 2d$

Proof:

We can compute the $\inf_{t<1/2} \dfrac{e^{-tx}}{(1-2t)^d}$ using 1^{st} order der.

Let $f(t) = \dfrac{e^{-tx}}{(1-2t)^d}$, $t < 1/2$, $x > 2d$.

then $f'(t) = \dfrac{-xe^{-tx}(1-2t)^d + 2de^{-tx}(1-2t)^{d-1}}{(1-2t)^{2d}}$

$= \dfrac{e^{-tx}(1-2t)^{d-1}}{(1-2t)^{2d}} [2d - x(1-2t)]$

Using that $x > 2d$, and for $t < 1/2$

inf is attained when $2d - x(1-2t) = 0$

$\Rightarrow t = \dfrac{x-2d}{2x} \Rightarrow \mathbb{P}(X > x) = \dfrac{e^{-\left(\frac{x-2d}{2}\right)}}{\left(\frac{2d}{x}\right)^d}$

$= \exp\left(d \log\left(\frac{x}{2d}\right) - \frac{x-2d}{2}\right)$ ☑

I explicitly asked to prove this fact...

(c) Using the already given fact that

$$\log(1+u) \leq u - \frac{1}{4}\min\{u, u^2\} \quad \forall u \geq 0$$

and the result from b)

$$\mathbb{P}(X-2d > z) = \exp\left(d\log\left(1+\frac{z}{2d}\right) - z/2\right).$$

Let $u = z/2d$ then $0 \leq \frac{z}{2d} \leq 1 \iff 0 \leq u \leq 1$ $\Rightarrow$ $\min\{u, u^2\} = u^2$

$$\Rightarrow \mathbb{P}(X-2d > z) \leq \exp\left[d\left\{1 - \frac{1}{4}\min\left(\frac{z}{2d}, \frac{z^2}{4d}\right)\right\} - z/2\right)$$

$$= \exp\left(d - \frac{d}{4}\min\left(\frac{z}{2d}, z^2/4d\right) - z/2\right)$$

it follows that:

$$\Rightarrow \mathbb{P}(X-2d > z) \leq \begin{cases} \exp\left(-z^2/16d\right) & 0 \leq z \leq 2d \\ \exp\left(-z/8\right) & \text{if } z > 2d \end{cases}$$

T₀

C

(ii) we can write $P(X - 2d > z) \leq \exp\left(-\min\left(\frac{z^2}{16d}, \frac{z}{8}\right)\right)$

as

$$\Rightarrow P(X - 2d \leq z) \geq 1 - \exp\left(-\min\left(\frac{z^2}{16d}, \frac{z}{8}\right)\right)$$

let $\delta = \exp\left(-\min\left(\frac{z^2}{16d}, \frac{z}{8}\right)\right).$

$$\Rightarrow \log\delta = -\min\left(\frac{z^2}{16d}, \frac{z}{8}\right)$$

$$\Rightarrow -\log\delta = \min\left(\frac{z^2}{16d}, \frac{z}{8}\right) \leq \frac{z^2}{16d} + \frac{z}{8}.$$

$$\Rightarrow \frac{z^2}{16d} + \frac{z}{8} \geq -\log\delta.$$

We need to bound z using roots ~~of~~ above,

4

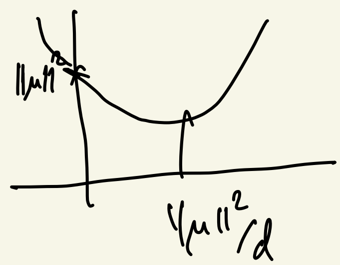
6.> <u>Stein's paradox</u>

(a) Risk function of a shrinkage estimator

$$R(\mu, s) = \mathbb{E}[\|sX - \mu\|]^2$$

$$= \mathbb{E}\left[\|s(X-\mu) + \mu(s-1)\|^2\right]$$

$$= \mathbb{E}\left[s^2 (X-\mu)^T (X-\mu) + \|\mu\|^2 (s-1)^2 + 2s(s-1)\mu^T (X-\mu)\right]$$



For $s < 0$,

risk increases with decreasing s, Hence $s = 0$ is dominating

$xsd - x\|\mu\|^2$

For $s > 1$, risk increases with increasing s; (we will also see this in the later part as we compute the minimum risk)

$\square$

(b)

let us minimize the expression $\circledast$

$$f(d) = \mathbb{E}\left[s^2 (x-\mu)^T(x-\mu) + \|\mu\|^2(s-1) + 2s(s-1)\mu^T(x-\mu)\right]$$

$$= s^2 d + \|\mu\|^2 (s-1)$$

Diff. wrt d

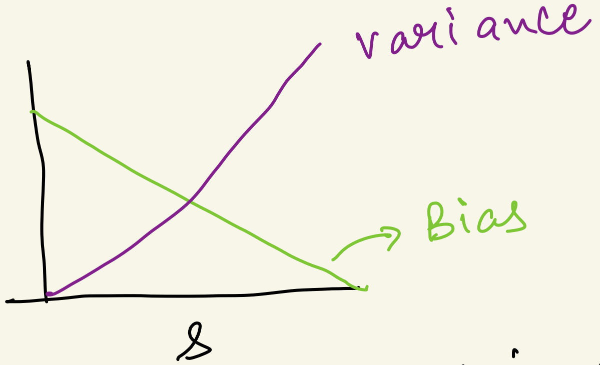
$$f'(d) = 2sd + 2(s-1)\|\mu\|^2 = 0$$

$$\Rightarrow s^* = \frac{\|\mu\|^2}{\|\mu\|^2 + d}$$

c) From the discussion we had during your office hour.

First $\hat{X} = sX$ is biased for $s \neq 1$ $(0 \leq s < 1)$

$\mathrm{var}(\hat{X}) = s^2 \, \mathrm{var}(X)$ decreases with decreasing s,



There is no bias-variance tradeoff:

$$E[\hat{\mu}^*] = s^* \mu \neq \mu$$

[the expectation of X]

The new estimator

$$\left(1 - \frac{d}{\|X\|^2}\right) X \cdot$$

considers shrinkage to be normalized by the choice of sample of X, as in the shrinkage is adjusted instead of being constant.

(d) Given the estimator $\hat{\mu}^{\delta} = \left(1 - \frac{\delta}{\|x\|^2}\right) x$

$$R(\delta) = R_{\mu}\left[\hat{\mu}^{\delta}\right] = \mathbb{E}\|\hat{x} - \mu\|^2$$

Let's centre the random variable & then minimize wrt δ.

$$= \mathbb{E}\|\hat{x} - x + x - \mu\|^2$$

$$= \mathbb{E}\left[\|\hat{x} - x\|^2 + \|x - \mu\|^2 + 2(\hat{x} - x)^T(x - \mu)\right]$$

$$= \mathbb{E}\left[\frac{\delta^2 \|x\|^2}{\|x\|^4}\right] + d + 2\mathbb{E}\left[(\hat{x} - x)^T(x - \mu)\right]$$

$$= \mathbb{E}\frac{\delta^2}{\|x\|^2} + d + 2\mathbb{E}\left[-\frac{\delta x^T(x - \mu)}{\|x\|^2}\right]$$

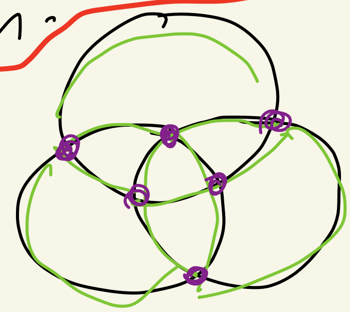
$$= \mathbb{E}\left[\frac{\delta^2}{\|x\|^2} - 2\frac{\delta x^T(x - \mu)}{\|x\|^2}\right] + d$$

$$= \mathbb{E}\left[\frac{\delta^2}{\|x\|^2}\right] - 2\delta\,\mathbb{E}\left[\frac{x^T(x - \mu}{\|x\|^2}\right] + d$$

7.) Planar Venn diagrams

Kudos!

I struggled with this one, & also referred to this article by Frank Ruskey, Carla .D. Savage called " The Search for simple symmetric venn diagrams" to get an idea of how venn diagram connects to a planar graph.



Proof:

let us say a planar circle venn diagram exists for $n \geq 5$.

Then One can construct vertices of Graph from the intersection of any two circles:

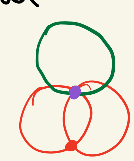
[circles intersect twice] in this case

What is C_2

$$V = 2 \cdot {}^{n}C_2 = n(n-1)$$

$$E = \underbrace{4 \cdot {}^{n}C_2}$$

∵ Each vertex is degree 4 as two circles intersection leads to 4 edges at the intersection pts:



The subtlety here is that there are no "collisions" with 3 or more circles intersecting at one point. Can you explain why?

faces = Total no. of regions :

Since each region corresponds to $\bigcap_{j \in J} A_j$

$J \subseteq \{1, 2, \ldots, n\}$

There are total $- 2^n$ subsets of J

$\therefore$ # faces $:= F = 2^n$:

Now the following should hold : $V - E + F = 2$

$LHS = V - E + F = 2\,{}^nC_2 - 4\,{}^nC_2 + 2^n$

I don't see why this holds (algebra)

$= 2^n - 2 \cdot {}^nC_2 = 2^n - n(n-1)$

exponential Quadratic

for $n \geq 5$, $2^n - n(n-1) > 2$

can be shown using Calculus : $\boxtimes$.

<u>Note</u>, This proof seems to work for $n = 4$ as well

(Seems to be correct, modulo the subtlety above + presentation

+.