Lea Fakhir
HW 1 8803.

Total: 75/100 (B)

1)

a) 1st of all the MGF: $M_X(\lambda) = E[e^{\lambda X}]$ of X.

→ ∴ $M_X(\lambda) = E[e^{\lambda X}]$. → Multiply both sides by $e^{-\lambda 4}$.

$$\to M_X(\lambda)\, e^{-\lambda 4} = E[e^{\lambda X - \lambda 4}].$$

→ use taylor series expansion of e^x: and substitute to equation.

$$e^{\lambda Y} = \sum_{4=0}^{\infty} \frac{\lambda^4}{4!} x^4.$$

taking expectation:

$$M_X(\lambda) = E\left(\sum_{4=0}^{\infty} \frac{\lambda^4}{4!} x^4\right)$$

$$M_X(\lambda) = E(x^4) \sum_{4=0}^{\infty} \frac{x^4}{4!}$$

→ Multiply both sides by $e^{-\lambda 4}$:

$$M_X(\lambda)\, e^{-\lambda 4} = E(x^4) \sum_{4=0}^{\infty} \frac{x^4 e^{-\lambda 4}}{4!}$$

I'm lost where's n and where's 4.

→ choose an optimal λ → $\lambda = 4/4$. for some integer 4.

$$M_X(4/4)\, e^{-(4/4)4} = e^{-4} \sum_{j=0}^{\infty} \frac{(4/4)^j}{j!} E(x^j).$$

→ focus on dominent term:

$$M_X(4/4)\, e^{-(4/4)4} \geq e^{-4} \frac{(4/4)^4}{4!} E(x^4).$$

→ using Stirlings approximation $4! \sim \sqrt{2\pi 4}\,(4/e)^4$ we get:

$$\frac{(u/u)^u}{u!} \approx u^{-u}.$$

no, this is not what we get from Stirling...

$$\therefore \quad M_X(u/u)\, e^{-(u/u)u} \geq E(X^u)\, u^{-u}.$$

(Also, the inequality to prove is exact, not asymptotic.

→ take infimum over $\lambda > 0$ on LHS, over $u \in \mathbb{Z}_+$ on RHS:

$$\boxed{\inf_{\lambda > 0} M_X(\lambda)\, e^{-\lambda u} \geq \inf_{u \in \mathbb{Z}^+} E[X^u]\, u^{-u}.}$$

b) showing: $\inf_{\lambda>0} M_X(\lambda) e^{-\lambda u} \geq \frac{1}{2} \inf_{u \in \mathbb{Z}} E(X^{2u}) u^{-2u}$. → given X symmetric

→ since we are given X is symmetric, this means → $E[e^{\lambda x}] = E[e^{-\lambda x}]$.

$$\hookrightarrow \quad M_X(\lambda) = E\left[\frac{e^{\lambda x} + e^{-\lambda x}}{2}\right].$$

where $e^{\lambda x} + e^{-\lambda x}$ → from power series:

$$= 2 \sum_{u=0}^{\infty} \frac{\lambda^{2u} x^{2u}}{(2u)!}$$

$$\therefore \quad M_X(\lambda) = \sum_{u=0}^{\infty} \frac{\lambda^{2u}}{2u!} E[\cancel{X}^{2u}].$$

→ Now showing the proof:

from Markov's inequality, where $P(Y \geq a) \leq \frac{E[Y]}{a}$.

$$\therefore \quad P(X \geq u) = P(e^{\lambda x} \geq e^{\lambda u}).$$

$\rightarrow P(X \geq u) \leq \frac{E[e^{\lambda X}]}{e^{\lambda u}} = \frac{M_X(\lambda)}{e^{-\lambda u}}$.

$\rightarrow$ Since X is symmetric we have:

$$P(X \leq -u) = P(-X \geq u) \leq M_X(\lambda)\, e^{-\lambda u}.$$

By union bound: $P(|X| \geq u) = P(X \geq u) + P(X \leq -u)$.

$$\therefore P(|X| \geq u) \leq 2 M_X(\lambda)\, e^{-\lambda u}.$$ $\rightarrow$ divide by 2.

$$\frac{1}{2} P(|X| \geq u) \leq M_X(\lambda)\, e^{-\lambda u}. \quad (1).$$

$\rightarrow$ Applying moment bound:

$\rightarrow$ Now applying Markov's inequality to X^{2q}.

$$P(|X| \geq u) = P(X^{2q} \geq u^{2q}) \leq \frac{E[X^{2q}]}{u^{2q}}$$

thus: $\frac{1}{2} P(|X| \geq u) \leq \frac{1}{2} \frac{E[X^{2q}]}{u^{2q}}$.

Switched ~

from proof (1).

$$M_X(\lambda)\, e^{-\lambda u} \geq \frac{1}{2} \frac{E[X^{2q}]}{u^{2q}}.$$

$\rightarrow$ Take infimum over u on RHS and $\lambda > 0$ on LHS.

$$\boxed{\inf_{\lambda > 0} M_X(\lambda)\, e^{-\lambda u} \geq \frac{1}{2} \inf_{u \in \mathbb{Z}^+} E[X^{2q}]\, u^{-2q}}$$

$-$

2) to prove $u_X(t)$ is convex function $\to u_X(t) = \log E[e^{tX}]$

applying Young's inequality:

given: $|a^T b| \leq \|a\|_p \|b\|_q$. Assuming $X \to$ discrete distribution.

$\to tX = (t \cdot X) \leq \|t\| \, \|X\|$.

$\to \boxed{E[e^{tX}] \leq E[e^{\|t\|_p \|X\|_q}]}$

t and X are numbers, so $\|t\|_p = |t|$, $\|X\|_q = |X|$ for any p, q. So, this is just an identity (not producing anything).

take log:

$u_X(t) = \log E[e^{tX}] \leq \log E[e^{\|t\|_p \|X\|_q}]$

$\to$ this shows $u_X(t)$ convex.

or we can show convexity by taking second derivative: $f''(t) > 0$

so for any $\alpha \in [0,1]$ and any t_1, t_2 we show:

$u_X(\alpha t_1 + (1-\alpha) t_2) \leq \alpha u_X(t_1) + (1-\alpha) u_X(t_2)$

$\to$ proving this shows $u_X(t)$ convex.

$\to$ to prove convexity we will establish that the MGF:

$M_X(t) = E[e^{tX}]$.

log convex means:

$M_X(\alpha t_1 + (1-\alpha) t_2) \leq M_X(t_1)^\alpha \, M_X(t_2)^{1-\alpha}$.

taking logarithms $u_X(t)$:

using Young's inequality in the form of Holder inequality.

for p and q (i.e. $\frac{1}{p} + \frac{1}{q} = 1$).

$$\boxed{E[e^{\alpha t_1 X + (1-\alpha) t_2 X}] \leq E[e^{t_1 X}]^\alpha \, E[e^{t_2 X}]^{1-\alpha}}$$

take logarithm:

$u_X(\alpha t_1 + (1-\alpha) t_2) = \log M_X(\alpha t_1 + (1-\alpha) t_1)$

this is (explicitly) Young's inequality in the form

Rather, it's the one with $\|\cdot\|_{L_p}$ and $\|\cdot\|_{L_q}$ norm... I wrote it.

$$\leq \log\left(M_X(t_1)^{\alpha} M_X(t_2)^{1-\alpha}\right).$$

using logarithm properties:

$$\log A^{\alpha} B^{1-\alpha} = \alpha \log A + (1-\alpha)\log B.$$

we conclude.

$$u_X(\alpha t_1 + (1-\alpha)t_2) \leq \alpha\, u_X(t_1) + (1-\alpha) u_X(t_2).$$

→ thus $u_X(t)$ is convex.

+/-

3) Proving the following upper and lower bounds → on the Gaussian tail probability:

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

where → $\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$ → standard normal PDF.

$\Phi(u) = \int_u^{\infty} \phi(t)\,dt$ is the tail probability of the standard normal distribution.

→ proving upper bound:

$$\Phi(u) \leq \frac{1}{u}\phi(u).$$

Use integration by parts: set:

* $f(t) = \Phi(t)$ so that $f'(t) = \phi(t)$.
* choose $g'(t) = 1$ so $g(t) = t$

$$\rightarrow \int_u^{\infty} \phi(t)\,dt = \left[-\frac{\phi(t)}{t}\right]_u^{\infty} + \int_u^{\infty} \frac{\phi(t)}{-2}\,dt.$$

Since $\phi(t) \to 0$ as $t \to \infty$, so the boundary term at ∞ vanishes:

$$\Phi(u) = \frac{\phi(u)}{} + \int^{\infty} \phi(t)\,dt.$$

$4 \qquad 4 \quad \overline{t^2}$

→ Since $\int_4^\infty \frac{\phi(t)}{t^2} dt \geq 0 \rightarrow$ we get the upper bound:

$$\boxed{\phi(4) \leq \frac{1}{4}\phi(4).}$$

+

Now proving lower bound:
from the 1st part we have:

$$\phi(4) = \frac{\phi(4)}{4} + \underbrace{\int_4^\infty \frac{\phi(t)}{t^3} dt.}_{\left(\frac{-\phi t}{t^3}\right)_4^\infty + \int_4^\infty \frac{3\phi d t}{t^4}}$$

∴

$$\phi(4) = \frac{\phi(4)}{4} - \frac{\phi 4}{4^3} + \underbrace{\int_4^{\infty} \frac{3\phi(t)}{t^4} dt.}_{\text{bigger than } 0.}$$

(a) ~~t~~..).

$$\therefore \boxed{\phi(4)\left(\frac{1}{4} - \frac{1}{4^3}\right) < \phi(4)}$$

+

Thus lower and upper bound proven.

$$\therefore \left(\frac{1}{4} - \frac{1}{4^3}\right)\phi(4) \leq \phi(4) \leq \frac{1}{4}\phi(4)$$

b) The integration by parts used in part a: $\phi(4) = \frac{\phi(4)}{4} + \int_4^\infty \frac{\phi(t)}{t^2}\,dt$.

now here we don't drop integral:

apply: $f(t) = \phi(t) \quad f'(t) = \phi'(t)$.

$g'(t) = \frac{1}{t^2} \quad g(t) = \frac{-1}{t}$.

∴ Applying integration by parts:

$$\int_4^\infty \frac{\phi(t)}{t^2}\,dt = \left[-\frac{\phi(t)}{t}\right]_4^\infty + \int_4^\infty \frac{\phi'(t)}{t^3}\,dt.$$

as $\phi(t) \to 0$ as $t \to \infty$ so that $t \to \infty$ vanishes.

$$\int_4^\infty \frac{\phi(t)}{t^2}\,dt = \frac{-\phi(4)}{4} + \int_4^\infty \frac{\phi'(t)}{t^3}\,dt.$$

now take equation from part A and:

→ applying to $\int_4^\infty \frac{3\phi(t)}{t^4}\,dt$ → below:

$$\int_4^\infty \frac{3\phi(t)}{t^5}\,dt = \left[-\frac{3\phi(t)}{t^5}\right]_4^\infty - \int_4^\infty \frac{(5\times3)\,\phi(t)}{t^6}\,dt.$$

+

$$= \frac{3\phi(4)}{4^5} - \int_4^\infty \frac{15\,\phi(t)}{t^6}\,dt.$$

→ where $-\int_4^\infty 15\,\phi(t)/t^6\,dt < 0$ → thus we find:

→ take as $t \to \infty$.

$$\therefore \phi(4) = \phi(4)\left(\frac{1}{4} - \frac{1}{4^3} + \frac{3}{4^5}\right)\phi(4).$$

+

Good Job!

3-2) . 1st take $\phi(u)$ as CDF of standard normal distribution.

$$\Phi(u) = \int_{-\infty} \phi(t)\,dt \quad \text{with} \quad \phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

→ substitute the standard normal density function.

$$\frac{1}{2} - \Phi(u) = \frac{1}{\sqrt{2\pi}} \int_u^{\infty} e^{-t^2/2}\,dt. \quad \ldots\ldots \quad (1)$$

→ Now to get the power series expansion we'll use the integration by parts.

→ change variable: $t^2 = u^2 + s$.

Differentiating: $2t\,dt = ds \rightarrow dt = \frac{ds}{2t}$.

Rewrite integral: in terms of s: from equation (1):

where

$$\int_0^{\infty} e^{-t^2/2}\,dt = \int_0^{\infty} e^{-(u^2+s)/2} \frac{ds}{2(u+s)}. \quad \ldots \quad (2)$$

Use taylor series expansion on $e^{-s/2}$: $t \neq u+s$

$$e^{-s/2} = \sum_{u=0}^{\infty} \frac{(-s/2)^u}{u!}$$

∴ plugging back: to equation (2):

$$\int_0^{\infty} e^{-u^2/2} \sum_{u=0}^{\infty} \frac{(-s/2)^u}{u!} \frac{ds}{2(u+s)}.$$

$$\sum_{u=0}^{\infty} \frac{(-1)^u}{2^u u!} e^{-u^2/2} \underbrace{\int_0^{\infty} \frac{s^u}{2(u+s)}\,ds.}$$

✓ using gamma function identity.

A) follows the gamma distribution $\int_0^\infty \frac{x^u}{1+x}\,dx = \boxed{\frac{\pi}{\sin((u+1)\pi)} \neq \frac{\pi}{(u+1)\pi}} = \frac{1}{u+1}.$

we get: $\therefore \int_0^\infty \frac{s^u}{2(u+s)}\,ds = \frac{e^{u+1}}{2(u+1)}.$

$\therefore$ Plugging back:

$$\sum_{u=0}^{\infty} \frac{(-1)^u}{2^u u!} e^{-u^2/2} \frac{u^{2u+1}}{2(2u+1)}.$$

→ since each term in summation has $e^{-u^2/2}$ factor it out and rewrite in this format.

$$\therefore \frac{1}{\sqrt{2\pi}} \sum_{u=0}^{\infty} \frac{(-1)^u \, u^{2u+1}}{2^u u! \, (2u+1)}$$

→ Now plug back to equation ①:

$$\boxed{\frac{1}{2} - \phi(u) = \frac{1}{\sqrt{2\pi}} \sum_{u=0}^{\infty} \frac{(-1)^u \, u^{2u+1}}{2^u u! \, (2u+1)}.}$$

4) i) Define an indicator function:

$I = 1(X \geq (1-t)E(X))$

Multiply both sides by X and take expectation

$E[XI] \geq (1-t)E[X]\, P(X \geq (1-t)E(X))$.

$\to$ since I is 1 whenever $X \geq (1-t)E(X)$ so in that region $XI \geq (1-t)E(X)I$. and taking expectation maintains the inequality:

$$\to \quad E[XI] \geq (1-t)E(X)\, P(X \geq (1-t)E(X)).$$

apply Cauchy Schwarz: where:

$$E[XI] \leq \sqrt{E(X^2)E(I^2)}.$$ since $I^2 = I$ (indicator function values in $\{0,1\}$

$$E[I^2] = E[I] = P(X \geq (1-t)E(X)$$

Thus,

$$E[XI] \leq \sqrt{E(X^2)P(X \geq (1-t)E(X)}$$ $\to$ solve for $P(X \geq 1-t\,E(X)$

Lower bound from step ①:

$$E[XI] \geq (1-t)E(X)\, P(X \geq (1-t)E(X)).$$

upper bound from Cauchy-Schwarz:

$$E(XI) \leq \sqrt{E(X^2)P(X \geq (1-t)E(X)}.$$

Use both bounds:

$$(1-t)E(X)P(X \geq (1-t)E(X)) \leq \sqrt{E(X^2)P(X \geq (1-t)E(X)}$$

divide by $\sqrt{P(X \geq (1-t)E(X))}$ and then square:

$$\to \quad (1-t)^2 E(X)^2 P(X \geq (1-t)E(X) \leq E(X^2))$$

→ Now solve for $p(x \geq 1-t\,(E(x)))$.

$$p(x \geq (1-t)\,E(x)) \geq \frac{(1-t)^2\,E(x)^2}{E(x^2)}$$

→ Now substitute $1-t$ by t:

$$\boxed{p(x \geq (1-t)\,E(x)) \geq \frac{t^2\,E(x)^2}{E(x^2)}}$$

+ Correct!

ii) Recall variance of x:

$Var(x) = E(x^2) - E(x)^2$

→ Now the probability bound: from Paley-Zygmund inequality.

$$P(X \geq (1-t)E(X)) \geq t^2 \frac{E(x)^2}{E(x^2)}.$$

→ to strengthen the bound → introduce Var(x) into denom:

$$E(x^2) = Var(x) + E(x)^2.$$

$$\therefore \frac{E(x^2)}{E(x^2)} = \frac{E(x)^2}{Var(x) + E(x)^2}.$$

→ incorporate t^2 since strengthened inequality modifies the denominator to include $t^2 E(x)^2$ → we want denom to account for variability in x and scaling by t^2 → to ensure the bound becomes sharper.

$$\therefore \boxed{P(X \geq (1-t)E(x)) \geq t^2 \frac{E(x)^2}{t^2 E(x^2) + Var(x)}}$$

→ improves the Paley-Zygmund bound.

An example of the sharpness: 2-point RV:

→ Let X be a random variable that takes 2 distinct values a and b with prob p and 1-p.

$$X = \begin{cases} a & \text{prob } p \\ b & \text{prob } 1-p. \end{cases}$$

But I see no derivation of Cantelli ...

$\therefore E(X) = pa + (1-p)b.$

$E(X^2) = pa^2 + (1-p)b^2.$

$\therefore Var(X) = pa^2 + (1-p)b^2 - (pa + (1-p)b)^2.$

→ check probability

$P(X \geq (1-t)E(X))$ → the threshold value is $X \geq (1-t)E(X)$

$\therefore$ 1) if $b \geq (1-t)E(X)$, then the prob → $P(X \geq (1-t)E(X)) = 1-p.$

2) if $a < 1-t(E(X)$ → the prob is 0.

extra.

$\therefore P(X \geq (1-t)E(X)) = 1-p$ if $b \geq (1-t)E(X)$

where $a=0$ $b=1$ and $p=\frac{1}{2}$ (X is bernoulli $\frac{1}{2}$).

then $E(X) = \boxed{\frac{1}{2}}$

$Var(X) = \frac{1}{2} - \frac{1}{4} = \boxed{\frac{1}{4}}$

$Var(X) = p(1-p) = \frac{1}{4}$, yes.

→ the bound becomes:

$$P(X \geq (1-t)E(X)) \geq \frac{t^2 (\frac{1}{2})^2}{t^2(\frac{1}{2})^2 + \frac{1}{4}}$$

$$= \boxed{\frac{t^2}{t^2+1}}$$

$LHS = \mathbb{P}\{X \geq (1-t)\mathbb{E}[X]\}$

$= \mathbb{P}\{X \geq (1-t)\frac{1}{2}\} = \frac{1}{2}$

$= \mathbb{P}\{X = 1\}$ $\forall t \in (0,1)$

$RHS = \frac{t^2}{t^2+1} < \frac{1}{2}$ (strictly) $\forall t \in (0,1)$.

Check:

for $t=1$ → the inequality becomes:

$P(X \geq 0) = 1$ → $\frac{1^2}{1^2+1} = \frac{1}{2}$.

this matches the actual probability (Bernoulli ($\frac{1}{2}$)) which demonstrates equality for this bernoulli example, thus Cantelli's inequality is sharp in this case

So, this example shows tightness of Cantelli only for $t=1$.
(Note that for $t=1$, Cantelli does not improve over P.-Z.!)

c) we can now assume $E[|X|^p] < \infty$ for $p>1$ and generalize the Paley-Zygmund inequality.

→ 1st apply Hölder's inequality:

it states that for $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ we have:

$$E[UV] \leq E[|U|^p]^{1/p} \cdot E[|V|^q]^{1/q}$$

take $U = XI$
$V = I$

we get:

$$E(XI) \leq E[|X|^p]^{1/p} \, E[(I)^q]^{1/q}$$

Since $I^q = I$, simplify:

$$E[XI] \leq E[|X|^p]^{1/p} \, P(X \geq (1-t) E(X))^{1/q}$$

We lower bound:

$$E[XI] \geq (1-t) E(X) \, P(X \geq (1-t) E(X))$$

we get:

$$(1-t) E(X) P(X \geq (1-t) E(X)) \leq E(|X|^p)^{1/p} \, P(X \geq (1-t) E(X))^{1/q}$$

solve for $P(X \geq (1-t) E(X))$:

$$P(X \geq (1-t) E(X))^{1-\frac{1}{q}} \leq \frac{E(|X|^p)^{1/p}}{(1-t)(E(X)}$$

We $q = \frac{p}{p-1}$ so $1 - \frac{1}{q} = \frac{1}{p-1}$ + we raise to the power $p-1$

$$\boxed{P(X \geq 1-t(E(X)) \geq \left(\frac{t^p \, E[|X|^p]}{E(|X|)^p}\right)^{1/p-1}}$$

(+) Well done!

5)

a) given that $X \sim \chi^2_{2d}$ → chi squared distribution with 2d deg freedom.

→ the MGF: for chi squared distributed variable: $X \sim \chi_u^2$ with u degrees of freedom is given by:

MGF defined as:

$M_Z(t) = E[e^{tx}]$.

where $X = Z_1^2 + Z_2^2$ and $Z_1, Z_2 \sim N(0,1)$ (indep standard normal RV's).

$\therefore M_Z(t) = E[e^{t(Z_1^2 + Z_2^2)}]$ → using indep

$M_Z(t) = E[e^{tZ_1^2}] \, E(e^{tZ_2^2})$

→ this reduces the problem to computing the MGF of a single-chi squared variable with 1 degree of freedom.

$\therefore$ finding $M_Y(t)$ for a single Z^2 term:

Define $Y = Z^2$ $\quad Z \sim N(0,1)$

$\therefore$ PDF is: $f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$, $y > 0$

$$\therefore M_Y(t) = E[e^{tY}] = \int_0^\infty e^{ty} \frac{1}{\sqrt{2\pi y}} e^{-y/2} dy$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi y}} e^{-(1/2 - t)y} dy$$

use Gamma function identity:

$$\int_0^\infty y^{\alpha-1} e^{-By} dy = \frac{\Gamma(\alpha)}{B^{(\alpha)}} \quad \text{for } \alpha > 0, B > 0$$

→ in our case $\alpha = 1/2$ $\quad B = 1/2 - t$

$\therefore M_Y(t) = (1-2t)^{-1/2}$ $\quad t < \frac{1}{2}$.

Why not to use that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1$$

instead of some obscure identity that one googled anyway? (^_^)

$\therefore$ for $M_Z(t) = (1-2t)^{-1/2} \times (1-2t)^{-1/2}$

$= (1-2t)^{-1}$.

$\therefore$ $\boxed{M_Z(t) = \frac{1}{1-2t} \quad t<\frac{1}{2}}$.

$\therefore$ → since $X = \sum_{i=1}^{2d} Z_i^2$, the mgf of X is: $M_{2d}(t) = E[e^{t(Z_1^2+Z_2^2 \dots Z_{2d}^2)}]$

which becomes: by independence

$\Rightarrow M_{2d}(t) = \prod_{i=1}^{2d} M_{Z_i^2}(t)$

where $M_{Z_i^2} = (1-2t)^{-\frac{1}{2}} \quad t<\frac{1}{2}$.

$\therefore$ $\boxed{M_{2d}(t) = (1-2t)^{-\frac{2d}{2}} = (1-2t)^{-d} \quad t<\frac{1}{2}}$

+0

b) Chernoff's method states that for any $t>0$

$P(X>x) = P(e^{tX} > e^{tx}) \le E[e^{tX}]e^{-tx}$.

→ replace MGF $X \sim \chi^2_{2d}$

$P(X>x) \le \inf_{t<\frac{1}{2}} e^{-tx} M_{2d}(t)$.

$P(X>x) \le \inf_{t<\frac{1}{2}} e^{-tx} (1-2t)^{-d}$

→ take the logarithm

$\log P(X>x) \le \inf_{t<\frac{1}{2}} (-tx - d\log(1-2t))$

now differentiate both sides:

$\frac{d}{dt}(-tx - d\log(1-2t)) = -x + \frac{2d}{1-2t}$.

→ set derivative $=0$ → to find critical points, find minimum

$-x + \frac{2d}{1-2t} = 0$

$$\frac{2d}{x} = 1-2t$$

$$t = \frac{1}{2}\left(1 - \frac{2d}{x}\right)$$

→ now substitute optimal t back:

$$P(X > x) \geq \exp\left(d\log\left(\frac{x}{2d}\right) - \frac{x-2d}{2}\right)$$

where $z = x - 2d$.

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

→ this shows the probability of deviations from the mean decays exponentially

+

c) we found in part B:

$P(X - 2d > z) = \exp\left(d\log\left(\frac{2d+z}{2d}\right) - \frac{z}{2}\right)$ for $0 \leq z \leq 2d$ → exponential quadratic

$z > 2d$ → expon linear decay

→ use inequality given in hint:

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

– I asked to prove this inequality first.

set $u = \frac{z}{2d}$:

$\log\left(\frac{2d+z}{2d}\right) = \log\left(1 + \frac{z}{2d}\right)$ → apply inequality

+/–

$$\log\left(1+\frac{z}{2d}\right) \leq \frac{z}{2d} - \frac{1}{4}\min\left\{\frac{z}{2d}, \left(\frac{z}{2d}\right)^2\right\}$$

→ substitute the upper bound for $\log\left(1+\frac{z}{2d}\right)$:

$$\leq \exp\left(d\left(\frac{z}{2d} - \frac{1}{4}\min\left\{\frac{z}{2d}, \left(\frac{z}{2d}\right)^2\right\} - \frac{z}{2}\right)\right.$$

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

Now we can analyze 2 cases:

for the case: 1 : $0 \le z \le 2d$.

here $\frac{z}{2d} \le 1$ so minimum is $\left(\frac{z}{2d}\right)^2$:

$$p(x - 2d > z) \le \exp\left(\frac{-d}{4} \cdot \frac{z^2}{4d^2}\right)$$

$$= \exp\left(\frac{-z^2}{16d}\right).$$

Case 2 : $z > 2d$

Here $\frac{z}{2d} > 1$ so minimum is $\frac{z}{2d}$:

$$p(x - 2d > z) \le \exp\left(\frac{-d}{4} \cdot \frac{z}{2d}\right)$$

$$= \exp\left(\frac{-z}{8}\right)$$

$\therefore$ we conclude: $p(x - 2d > z) \le \begin{cases} \exp\left(-\frac{z^2}{16d}\right) & 0 \le z \le 2d \\ \exp\left(-\frac{z}{8}\right) & z > 2d \end{cases}$

$\therefore$ we rewrite:

$$\boxed{p(x - 2d > z) \le \exp\left(-\min\left(\frac{z^2}{16d}, \frac{z}{8}\right)\right)}$$

$\rightarrow$ solving for z: $\rightarrow$ $\min\left(\frac{z^2}{16d}, \frac{z}{7}\right) = \log\frac{1}{\delta}$

$z = \sqrt{16d \log\frac{1}{8}}$ $\quad 0 < z \le 2d$

$z = 8 \log\frac{1}{\delta}$ $\quad z > 2d$

Comparing: $x - 2d \le \sqrt{cd \log(1/\delta)} + c\log(1/\delta)$

$\therefore \boxed{c = 16 \qquad c = 8}$

$\therefore p(x - 2d > z) \le \begin{cases} \exp\left(-\frac{z^2}{16d}\right) & 0 < z \le 2d \\ \exp\left(-\frac{z}{8}\right), & z > 2d. \end{cases}$

→ c(ii).

as stated in given, let $P(X - 2d > z) = \delta$.

$\therefore \delta = \exp(-\min\{\frac{z^2}{16d}, \frac{z}{8}\}$.

→ take the logarithm and get rid of ⊖ sign:

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

for condition $0 \le z \le 2d \to -\log(\delta) = \frac{z^2}{16d}$.

$\therefore$ solving for $z \to z = \sqrt{16d \log\frac{1}{\delta}}$

→ for condition $z > 2d \to \frac{z}{8} = -\log\delta$

$\therefore \boxed{z = -8\log\delta}$

$\therefore z \le \max\{\sqrt{16d\log\frac{1}{\delta}}, 8\log\frac{1}{\delta})$

$\to z \le (\max\sqrt{16d\log\frac{1}{\delta}} + 8\log\frac{1}{\delta})$

+

6)

a) given: $\mu_S = sx$ → we will use:

1st from given and define:

$$\text{Risk}_\mu[\hat{y}] = E_y[\|\hat{y} - y\|^2].$$

→ for MLE estimator $\mu = x$.

$$\text{Risk}_y[x] = E_y[\|x - y\|^2]. \quad \dots (1).$$

Since $x \sim N(y, I_d)$ we decompose: $x = \mu + z \rightarrow z = x - y$.
→ where $z \sim N(0, I_d)$

→ thus replacing into eq (1):

$$\text{Risk}_y[x] = E_y[\|z\|^2] = E[\sum_{i=1}^{d} z_i^2] = d.$$

→ Use the shrinkage estimator we get: $\mu_S = sx$:

$$\text{Risk}_\mu[\mu_S] = E_\mu[\|sx - y\|^2] =$$

$$\|sx - y\|^2 = \|s(y + z) - y\|^2 = \|sy + sz - y\|^2$$
$$= \|(s-1)y + sz\|^2.$$

→ Now power squared can be rewritten:

$$\|(s-1)y + sz\|^2 = \|(s-1)\mu\|^2 + \|sz\|^2 + 2(\langle s-1\,\mu, sz\rangle). \quad (1)$$

→ Put expectation back:

$$E[\text{Risk}_y[\hat{y}_s]] = (s-1)^2 \|y\|^2 + s^2 E(\|z\|^2).$$

since $E[\|z\|^2] = d$.

$\rightarrow$ $\boxed{\text{Risk}_y[\psi_s] = (s-1)^2 \|y\|^2 + s^2 d.}$ +

$\rightarrow$ to dominate the MLE $(s=1)$ we need:

$\text{Risk}_y(x)$.

$$(s-1)^2 \|y\|^2 + s^2 d < d.$$

$$\therefore (s-1)^2 \|y\|^2 < d(1-s^2).$$

$$\frac{(s-1)^2}{s^2-1} < \frac{1}{\|y\|^2}.$$

for large $\|y\|^2$ we can approximate.

$$s = 1 - \frac{c}{\|y\|^2}.$$

where c is a small constant. This means some shrinkage for $0 < s < 1$ dominates the MLE.

$\rightarrow$ for $s > 1$ $\rightarrow$ the risk increases, and is dominated by MLE $\hat{\psi} = x$.

$\rightarrow$ for $s < 0$

the risk formula ① becomes:

$$\text{Risk}_y[\psi_s] = \underbrace{s^2(\|y\|^2 + d)}_{\text{always } +} \underbrace{-2s\|y\|^2}_{\text{when } s<0 + \text{ becomes large increasing risk.}} + \|y\|^2.$$

Thus risk increases for $s < 0$ and $s > 1$. +

and thus attention should be focused on $(0,1)$ for shrinkage.

b) found from part a the quadratic risk function for the shrinkage estimator.

$\therefore$ we can now find optimal s^* by taking derivative,

$$\frac{d}{ds}\left((s-1)^2\|y\|^2 + s^2 d\right) = 2(s-1)\|y\|^2 + 2sd$$

$\to$ set to zero:

$$2(s-1)\|y\|^2 + 2sd = 0.$$

$$s^*(\|y\|^2 + d) = \|y\|^2,$$

$$s^* = \frac{\|y\|^2}{d+\|y\|^2}.$$

$$\therefore \boxed{s^* = 1 - \frac{d}{d+\|y\|^2},}$$

c) The estimator $\mu = s^* x = \left(1 - \frac{d}{d+\|y\|^2}\right)x$ is not a proper estimator since an estimator should be a function of observed data x only, but S^* depends on $\|y\|^2$ also, which is not known. → this means we can't use $\|y\|^2$ in constructing y^*.

Thus, using the $y = \left(1 - \frac{d}{\|x\|^2}\right) x$ ensures y is an estimator of x only, which makes it valid.

→ this estimator ensures that when x is far from the origin + large $\|x\|^2$, the correction term is small → shrinkage.
→ when x is close to origin → small $\|x\|^2$ there is strong shrinkage towards 0.
This adapts the amount of shrinkage dynamically.

(I would simply say that $\|X\|^2$ is an unbiased estimate of $\|\mu\|^2 + d$.) +

d) → from quadratic risk:

$$R_y(y) = E_y \|\hat{y} - y\|^2].$$

for shrinkage estimator:

$$\hat{y}^S = \left(1 - \frac{\delta}{\|x\|^2}\right) x.$$

$$\|y^S - y\|^2 = \left\| 1 - \frac{\delta}{\|x\|^2} x - y \right\|^2.$$

Since $x = y + z$ where $z \sim N(0, I_d)$ we substitute.

$$y^S = \left(1 - \frac{\delta}{\|x\|^2}\right) y + z.$$

$$\therefore \quad \|y^S - y\|^2 = \left\|\left(1 - \frac{\delta}{\|x\|^2}\right)y + \left(1 - \frac{\delta}{\|x\|^2}\right)z - y\right\|^2.$$

$$= \left\|\left(1 - \frac{\delta}{\|x\|^2}\right)y + \left(1 - \frac{\delta}{\|x\|^2}\right)z\right\|^2.$$

$\rightarrow$ expand the square:

$$\dagger \quad R_y(\delta) = E_y\left[\frac{\delta^2}{\|x\|^4}\|y\|^2 + \left(1 - \frac{\delta}{\|x\|^2}\right)^2\|z\|^2 + 2\frac{\delta}{\|x\|^2}\left(1 - \frac{\delta}{\|x\|^2}\right)\langle y, z\rangle\right]$$

use $E_y\|z\|^2 = d$, $E_y[\langle y, z\rangle] = 0 \rightarrow$

$$\dagger \quad R(y(\delta)) = E_y\left[\frac{\delta^2}{\|x\|^4}\|y\|^2 + \left(1 - \frac{\delta}{\|x\|^2}\right)^2 d\right].$$

+ use Stein's lemma: $E_y\left[\frac{1}{\|x\|^2}\right] = \frac{1}{d-2}$ for $d \geq 3$.

and differentiate $R(\delta)$ & and find wish minimized $= 0$.

$$\frac{d}{d\delta} R(\delta) = E\left[\frac{2\delta}{\|x\|^4}\|y\|^2 - 2\left(1 - \frac{\delta}{\|x\|^2}\right)\frac{d}{\|x\|^2}\right] = 0$$

$$E\left[2\delta\frac{(d + \|y\|^2)}{\|x\|^4}\right] = E\left[\frac{2d}{\|x\|^2}\right]$$

$\therefore \delta^* = d - 2.$

$\therefore$ from the James-Stein estimator:

$$\boxed{y^{JS} = \left(1 - \frac{d-2}{\|x\|^2}\right)x.}$$

$\dagger$

7) Eulers formula states that for any planar graph with:

v, e, f (regions into which $\mathbb{R}^2$ is divided)

$$v - e + f = 2.$$

→ our goal is to estimate v_n, e_n, f_n in terms of their values for $n-1$.

① Estimating # faces f_n.

A venn diagram for n sets should divide the plane into exactly 2^n regions, corresponding to all possible set interactions.

Thus: for a valid venn diagram: $f_n = 2^n$.

+

No need for this

~~Use recursion:~~
~~when $n=1$ → we have 2 regions so $f_1 = 2$.~~
~~when $n=2$ → the plane is divided into 4 regions~~
~~$f_2 = 2^2 = 4$.~~
~~f_3 = now doubles into $f_3 = 8$.~~

~~$\therefore$ $\boxed{f_n = 2 f_{n-1} = 2^n.}$~~

~~→ Additionally, each new curve adds at most $2(n-1)$ new regions~~

This is not obvious at all... (And if we had it just like that, we would immediately get a contradiction without Euler's...)

using Euler: for planar graph.

$v - e + f = 2.$

$f_n = 2 + e_n - v_n.$

$f_{n-1} = 2 + e_{n-1} - v_{n-1}. + 2(n-1).$

$\therefore f_n \leq f_{n-1} + 2(n-1).$

$\therefore$ we find $f_n \leq 2 \sum_{u=1}^{n-1} u = \frac{2(n-1)n}{2} = O(n^2).$

Now v_n:

→ Each new curve intersects exiting curves at most $2(n-1)$ times, this means the nth curve introduces at most $2(n-1)$ vertices.

→ trying to visualize.

one curve has $v_1 = 0.$

adding another → $v_2 \leq 2(1) = 2.$

$v_3 \leq v_2 + 2(2) = 2 + 4 = 6.$

$\therefore \boxed{v_n \leq v_{n-1} + 2(n-1).}$

→ expanding the recurrence:

$v_n \leq v_{n-2} + 2(n-2) + 2(n-1)$

$v_n \leq v_{n-3} + 2(n-3) + 2(n-2) + 2(n-1).$

$\therefore v_n \leq \sum_{u=1}^{n-1} 2u. = n(n-1).$

Now upper bound for E_n:

$$E_n \leq E_{n-1} + 2(n-1).$$

→ expanding:

$$E_n \leq E_{n-2} + 2(n-2) + 2(n-1)$$

$$E_n \leq 1 + \sum_{k=1}^{n-1} 2k.$$

$$E_n \leq 1 + \frac{2(n-1)n}{2}.$$

$$\therefore \boxed{E_n \leq 1 + (n-1)n.}$$

→ lower bound for E_n:

→ considering a venn diagram → each new curve creates $2(n-1)$ new edges → but if the venn diagram has higher n then every new curve will intersect in even more edges. → at least double:

$$\rightarrow \boxed{E_n \geq E_{n+1} + 4(n-1).}$$

Now, here using Euler's, you would conclude that $F_n = O(n^2)$.

→ the quadratic growth of E_n, F_n suggests that the # of intersections and faces rapidly increase making it impossible to construct a planar venn diagram for $n \geq 5$.

→ the lower and upper bounds are tight signifying that the estimates are asymptotically correct.

+/–