Stat135 HW 2

1) $\hat{\mu}_{m+n} = \alpha \bar{X}_n + \beta \bar{Y}_m$

a For the estimator to be unbiased, $E(\hat{\mu}_{m+n}) = \mu$ $\qquad E(\bar{X}_n) = E(\bar{Y}_m) = \mu$

$$E(\hat{\mu}_{m+n}) = E(\alpha \bar{X}_n + \beta \bar{Y}_m) = \alpha E(\bar{X}_n) + \beta E(\bar{Y}_m) = \alpha\mu + \beta\mu = (\alpha+\beta)\mu$$

$\therefore$ since $(\alpha+\beta)\mu = \mu$, $\alpha + \beta = 1$.

b $var(\hat{\mu}_{m+n}) = var(\alpha \bar{X}_n + \beta \bar{Y}_m) = \alpha^2 var(\bar{X}_n) + \beta^2 var(\bar{Y}_m)$

$$= \alpha^2\left(\frac{\sigma^2}{n}\right) + \beta^2\left(\frac{\sigma^2}{m}\right)$$

since $var(\bar{X}_n) = \frac{\sigma^2}{n}$, $var(\bar{Y}_m) = \frac{\sigma^2}{m}$

$\downarrow$

Given that the estimator is unbiased, $\alpha+\beta = 1$. $\quad = \alpha^2\left(\frac{\sigma^2}{n}\right) + (1-\alpha)^2\left(\frac{\sigma^2}{m}\right)$

$$= \sigma^2\left(\frac{\alpha^2}{n} + \frac{(1-\alpha)^2}{m}\right)$$

To find α that minimizes the function, we take the derivative of the variance function.

$$\frac{d}{d\alpha}\left[\sigma^2\left(\frac{\alpha^2}{n} + \frac{(1-\alpha)^2}{m}\right)\right] = \frac{d}{d\alpha}\left[\frac{\alpha^2}{n} + \frac{(1-\alpha)^2}{m}\right] = \frac{2\alpha}{n} + \frac{-2(1-\alpha)}{m}$$

$$\frac{2\alpha}{n}\ \frac{-2(1-\alpha)}{m} = 0$$

$$\Updownarrow$$

$$\frac{2\alpha m - 2n(1-\alpha)}{nm} = 0$$

$$\Updownarrow$$

$$2\alpha m - 2n(1-\alpha) = 0$$

$$2\alpha m + 2n\alpha = 2n$$

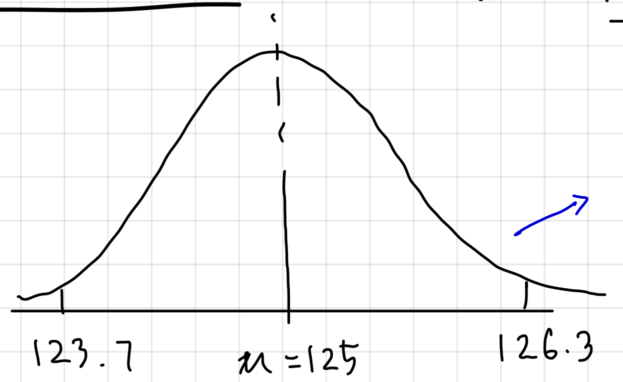
$$2\alpha(m+n) = 2n$$

$$\alpha = \frac{2n}{2(m+n)} = \boxed{\frac{n}{m+n}}$$

$$\beta = 1 - \alpha = 1 - \frac{n}{m+n} = \frac{m+n-n}{m+n} = \boxed{\frac{m}{m+n}}$$

Problem 2

$$\mu = \frac{123.7 + 126.3}{2} = 125$$



This is the CI for the 400 ppl sample.

99% CI

z score for 99% CI $= 2.6$

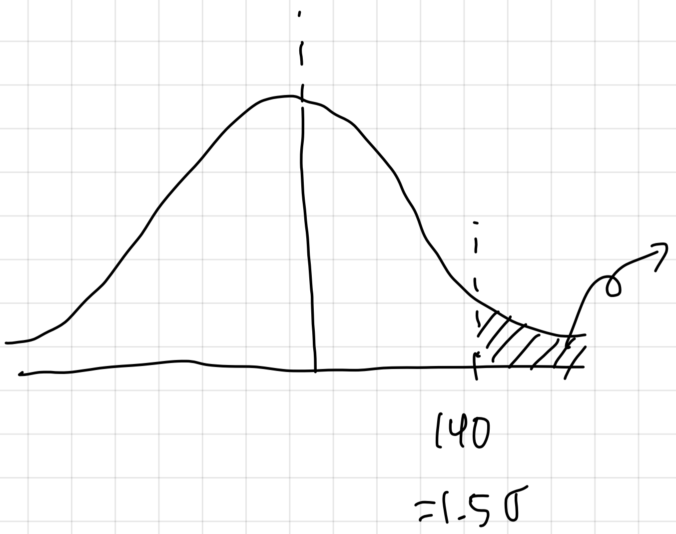
$$125 + 2.6\left(\frac{\sigma}{\sqrt{n}}\right) = 126.3$$

$$125 + 2.6\left(\frac{\sigma}{20}\right) = 126.3$$

$$2.6\left(\frac{\sigma}{20}\right) = 1.3$$

$$\frac{\sigma}{20} = \frac{1.3}{2.6} = 0.5$$

$$\sigma = 0.5 \times 20 = 10$$



patients with bp > 140mm,

$$1 - \Phi\left(\frac{140 - 125}{20}\right) = 1 - \Phi\left(\frac{15}{20}\right)$$

$$= 1 - \Phi(0.75)$$

$$\approx 1 - 0.933$$

$$\approx 0.067$$

95% CI for the proportion of population at risk:

$$0.067 \pm 1.96\sqrt{\frac{0.067(1-0.067)}{400}}$$

$$= [0.0425, 0.0915]$$

$$= [4.25\%, 9.15\%]$$

problem 3

a) 95% CI for μ.

$$\mu \pm 1.96 \frac{\sigma}{\sqrt{n}} = \mu \pm 1.96 \frac{(30)}{\sqrt{100}} = 85000 \pm \frac{1.96\ (\cancel{30}^{3})}{\cancel{10}} = 85000 \pm 5.88$$

$$= [84994.12,\ 85005.88]$$

b) i) CI approximates the true average not the average of 100 measurements.

we can compute the exact average of 100 measurements.

iii) There is no probability to whether the CI contains the exact height or not. It either contains or doesn't contain the true height.

iv) CI provides a range in which 95% of mean estimates (that would be then used for estimating the mean) will fall not all the possible measurements.

c) 99% $\rightarrow \alpha = 0.01$

$$z\left(\frac{0.01}{2}\right) s_{\bar{x}} = z(0.005)\ s_{\bar{x}} = 2.80\, s_{\bar{x}} = 1$$

$$s_{\bar{x}} = \frac{1}{2.80} = \frac{30}{\sqrt{n}}$$

$$\therefore \sqrt{n} = 2.80 \times 30$$

$$n = (2.80 \times 30)^2 = \boxed{7056}$$

problem 4

a) $F(-X) = 1 - F(X)$

$$F(-X) = p(Z \le -x) = p(-Z \le -x) = p(Z \ge x) = 1 - F(X) \quad \text{QED}$$

b) $F(z(\alpha)) = \alpha \iff p(Z \le z(\alpha)) = \alpha$

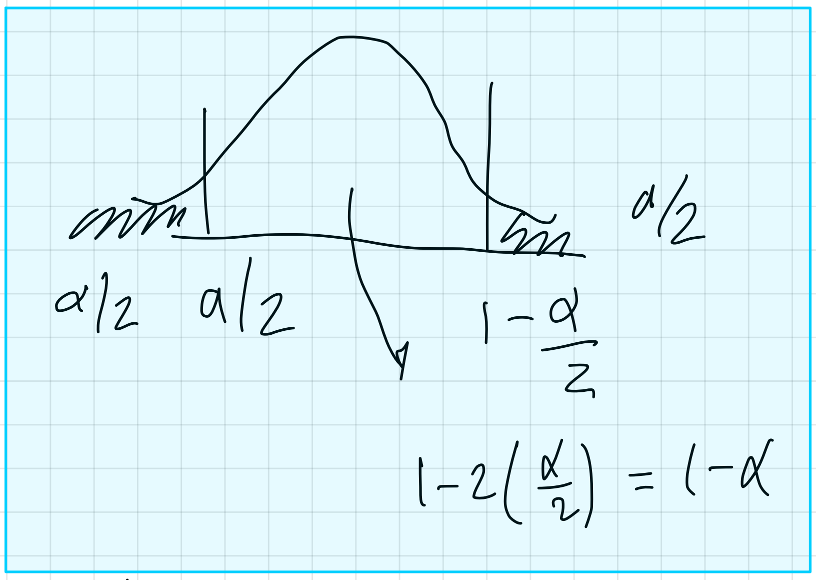
$F(z(1-\alpha)) = 1-\alpha \iff p(Z \le z(1-\alpha)) = 1-\alpha$

$p(Z \ge z(1-\alpha)) = \alpha$

then, $p(Z \le z(\alpha)) = p(Z \ge z(1-\alpha)) = p(-Z \ge z(1-\alpha))$

$= p(Z \le -z(1-\alpha))$

$\therefore \quad z(\alpha) = -z(1-\alpha) \quad$ QED



c) $F(b) - F(a) = 1-\alpha$

$p(a \le Z \le b) = 1-\alpha$

$$= \int_a^{z(\alpha/2)} f\,dx + \int_{z(\alpha/2)}^{z(1-\alpha/2)} f\,dx + \int_{z(1-\alpha/2)}^{b} f\,dx$$

$$= \int_a^{z(\alpha/2)} f\,dx + 1-\alpha + \int_{z(1-\alpha/2)}^{b} f\,dx = 1-\alpha$$

$$\Updownarrow$$

$$\int_a^{z(\alpha/2)} f\,dx + \int_{z(1-\alpha/2)}^{b} f\,dx = 0$$

What about the signs of $(b - z(1-\alpha/2))\,[a - z(\alpha/2)]$?

either one of the terms: $\int_a^{z(\alpha/2)} f\,dx$ or $\int_{z(1-\alpha/2)}^{b} f\,dx$ should be negative and the other should be positive for them to cancel out.

Since f is positive on $\mathbb{R}$,

Case #1

$$\int_{a}^{z(\alpha/2)} f\,dx > 0 \quad \text{if} \quad z(\alpha/2) > a.$$

$$\int_{z(1-\alpha/2)}^{b} f\,dx < 0 \quad \text{if} \quad z(1-\alpha/2) > b.$$

Case #2

$$\int_{a}^{z(\alpha/2)} f\,dx < 0 \quad \text{if} \quad z(\alpha/2) < a$$

$$\int_{z(1-\alpha/2)}^{b} f\,dx > 0 \quad \text{if} \quad z\left(1-\frac{\alpha}{2}\right) < b.$$

For case #1, the sign of $\left(b - z\left(1-\frac{\alpha}{2}\right)\right)\left(a - z\left(\frac{\alpha}{2}\right)\right)$ is positive

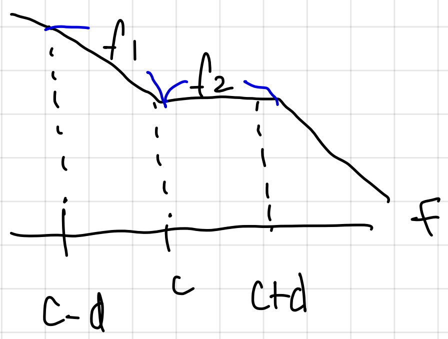
because $\ominus \times \ominus = \oplus$

For case #2, the sign of $\left(b - z\left(1-\frac{\alpha}{2}\right)\right)\left(a - z\left(\frac{\alpha}{2}\right)\right)$ is positive

because $\oplus \times \oplus = \oplus$

Hence, the sign of $\left(b - z\left(1-\frac{\alpha}{2}\right)\right)\left(a - z\left(\frac{\alpha}{2}\right)\right)$ is positive.

d) $\int_{c-d}^{c} f\,dx \geq \int_{c}^{c+d} f\,dx$

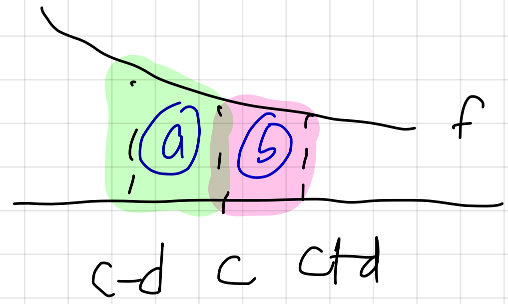


Let's say that the function f at $x \in [c-d, c]$ is f_1 and f at $x \in [c, c+d]$ is f_2.

Since $f_1 \geq f_2$ because f is a decreasing function, $f_1 \cdot |(c-d-c)| = cf_1 \geq f_2 \cdot |c+d-c| = cf_2$.

$\boxed{cf_1 \geq cf_2.}$ Hence, $\int_{c-d}^{c} f\,dx \geq \int_{c}^{c+d} f\,dx.$

Intuitive (graphical) interpretation is that since f is a decreasing function, the area ⓐ > ⓑ because we are integrating the same width (d) but by bigger f for $x \in [c-d, c]$ than the f for $x \in [c, c+d]$.



e) case 1) $h > z(1-\alpha/2)$

$e = z(\alpha/2) + t_1$

$h = z(1-\alpha/2) + t_2$

$h = z(\alpha/2) \quad e = z(1-\alpha/2)$

$$\int_{e}^{z(\alpha/2)} f\,dx + \int_{z(1-\alpha/2)}^{h} f\,dx = 0$$ from part c.

$$\int_{e}^{z(\alpha/2)} f\,dx = -\int_{z(1-\alpha/2)}^{h} f\,dx$$

$$\int_{e}^{z(\alpha/2)} f\,dx = \int_{z(1-\alpha/2)}^{z(1-\alpha/2)+t_2} f\,dx$$ ← based on the property that $z = -z$

$$\int_{z(\alpha/2)+t_1}^{z(\alpha/2)} f\,dx = \int_{z(1-\alpha/2)}^{z(1-\alpha/2)+t_2} f\,dx$$

$$\int_{-z(\alpha/2)-t_1}^{-z(\alpha/2)} f\,dx = \int_{z(1-\alpha/2)}^{z(1-\alpha/2)+t_2} f\,dx.$$

since $\int_{c-d}^{c} f\,dx \geq \int_{c}^{c+d} f\,dx$

$$\int_{-z(\alpha/2)-t_1}^{-z(\alpha/2)} f\,dx \geq \int_{-z(\frac{\alpha}{2})}^{-z(\frac{\alpha}{2})+t_1} f\,dx,$$

Hence, $-z(\alpha/2) = z(1-\alpha/2)$

For the equality to hold, $z(1-\frac{\alpha}{2}) + t_2 \geq -z(\frac{\alpha}{2}) + t_1$.

Therefore, $t_2 \geq t_1$. Hence, since $t_2 - t_1 \geq 0$, the interval that is the narrowest is $[z(\alpha/2), z(1-\alpha/2)]$ because t_2 and t_1 should both equal zero for the narrowest interval.

Problem 5

$$f(x;\sigma) = \frac{1}{2\sigma} e^{-|x|/\sigma}$$

$$E(X) = \int_{-\infty}^{\infty} \frac{x}{2\sigma} e^{-|x|/\sigma} dx = \int_{-\infty}^{0} \frac{-x}{2\sigma} e^{-x/\sigma} dx + \int_{0}^{\infty} \frac{x}{2\sigma} e^{-x/\sigma} dx$$

$$= \frac{1}{2\sigma}\left[\int_{-\infty}^{0} -x e^{-x/\sigma} dx + \int_0^{\infty} x e^{-x/\sigma} dx\right]$$

Since $E(X)=0$, we need to use the second moment.

these cancel out each other because the term on the right has positive x and the term on the left has negative x. But are the same magnitude.

$$E(X^2) = \int_{-\infty}^{\infty} \frac{x^2}{2\sigma} e^{-|x|/\sigma} dx = 2\int_0^{\infty} \frac{x^2}{2\sigma} e^{-|x|/\sigma} dx$$

$$\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt$$

$$\Gamma(3) = \int_0^{\infty} e^{-x} x^2 dx \Rightarrow \text{set } X = x/\sigma$$

$$dX = \boxed{\frac{1}{\sigma} dx}$$

$$= \int_0^{\infty} e^{-x/\sigma} \left(\frac{x}{\sigma}\right)^2 \frac{1}{\sigma} dx$$

$$\boxed{\int_0^{\infty} \left(\frac{1}{\sigma}\right)^3 x^2 e^{-\left(\frac{x}{\sigma}\right)} = \Gamma(3)}$$

$$\therefore E(X^2) = 2\int_0^{\infty} \frac{x^2}{2\sigma} e^{-|x|/\sigma} dx. = 2\sigma^2 \int_0^{\infty} \frac{x^2}{2\sigma^2 \cdot \sigma} e^{-|x|/\sigma} dx$$

$$= \frac{2\sigma^2}{2} \Gamma(3) = \frac{\cancel{2}\sigma^2(2)}{\cancel{2}}$$

$$= \boxed{2\sigma^2}$$

Since $E(X^2) = 2\sigma^2$, solving σ in terms of $\mu_2 = E(X^2)$ is

$$\sigma = \sqrt{\frac{1}{2}E(X^2)} = \sqrt{\frac{1}{2}\hat{\mu}_2} = \boxed{\sqrt{\frac{1}{2n}\sum_{i=1}^{n} x_i^2}}$$