ESE 542. hw 1. due. Sep 22, 2021

Problem 1.

a) For each peg, probability of guessing correctly is $P(\text{correct}) = \frac{1}{6}$

Therefore $P(\text{guessing all 4 pegs correctly}) = \frac{1}{6^4} = \frac{1}{1296}$

b) To put 3 red and 3 green blocks, firstly there's $\binom{6}{3}$ possibilities to choose positions for 3 red ones, then $\binom{3}{3}$ possibilities to choose the rest

Hence $\binom{6}{3}\binom{3}{3} = \frac{6!}{3!(6-3)!} \times 1 = 20$ patterns

Similarly for 9 blocks. there's

$$\binom{9}{3}\binom{6}{3}\binom{3}{3} = \frac{9!}{3!(9-3)!} \frac{6!}{3!(6-3)!} \times 1 = 1680 \text{ patterns}$$

Problem 2

a) $f(x|\alpha,\beta) = \frac{x^{\alpha-1} e^{-\frac{1}{\beta}x}}{\Gamma(\alpha)\beta^\alpha}$

$$= \frac{(\frac{1}{\beta})^\alpha x^{\alpha-1} e^{-\frac{1}{\beta}x}}{\Gamma(\alpha)}$$

let $t = \alpha - 1$

$$f(x|t,\beta) = \frac{(\frac{1}{\beta})^{t+1} x^t e^{-\frac{1}{\beta}x}}{\Gamma(t+1)}$$

Since. $\Gamma(t+1) = t\Gamma(t)$

$$f(x|t,\beta) = \frac{(\frac{1}{\beta}) \cdot (\frac{1}{\beta})^t x^t e^{-\frac{1}{\beta}x}}{t \cdot \Gamma(t)}$$

PDF property

Since $E(X) = \int_0^{+\infty} x f(x|\alpha,\beta)\,dx = \int_0^{+\infty} \frac{(\frac{1}{\beta})^\alpha x^\alpha}{\Gamma(\alpha)} e^{-\frac{1}{\beta}x}\,dx$ and $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$f(x|\alpha,\beta) = \int_0^{+\infty} \frac{(\frac{1}{\beta})(\frac{1}{\beta})^\alpha x^\alpha e^{-\frac{1}{\beta}x}}{\alpha \cdot \Gamma(\alpha)}\,dx = 1$$

$$\frac{\frac{1}{\beta}}{\alpha} E[X] = 1$$

$$E[X] = \alpha\beta$$

b). $Var(X) = E[(X-\mu)^2]$

$$= E[X^2] - (E[X])^2$$

$$E[X^2] = \int_0^{+\infty} x^2 f(x|\alpha,\beta)\,dx$$

$$= \int_0^{+\infty} \frac{(\frac{1}{\beta})^\alpha x^{\alpha+1} e^{-\frac{1}{\beta}x}}{\Gamma(\alpha)}\,dx$$

let $t = \frac{x}{\beta}$

$$E[X^2] = \int_0^{+\infty} \frac{(\frac{1}{\beta})^\alpha (\beta t)^{\alpha+1} e^{-t}}{\Gamma(\alpha)}\,d(\beta t)$$

$$= \int_0^{+\infty} \frac{(\frac{1}{\beta})^\alpha \beta^{\alpha+1}\cdot\beta\cdot t^{\alpha+1} e^{-t}}{\Gamma(\alpha)}\,dt$$

$$= \beta^2 \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha+1} e^{-t}\,dt \qquad \int t^{\alpha+1} e^{-t}\,dt = \Gamma(\alpha+2) \qquad (\alpha+1)!$$

$$= \beta^2(\alpha^2+\alpha)$$

$$\therefore Var(X) = \beta^2\alpha^2 + \beta^2\alpha - (\alpha\beta)^2$$

$$= \beta^2\alpha$$

c). for chi-squared distribution $X = \sum_{i=1}^{p} z_i^2,\ z_i \sim N(0,1)$

given $\chi_p^2 \sim \frac{x^{\frac{p}{2}-1} e^{-\frac{x}{2}}}{\Gamma(\frac{p}{2}) 2^{\frac{p}{2}}}$

for Gamma distribution $f(x|\alpha,\beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}$

therefore $\alpha = \frac{p}{2}$ $\quad \beta = 2$

Problem 3

for $X \sim Bernoulli(p)$, $\quad P(X=x) = \begin{cases} 1-p, & x=0 \\ p, & x=1 \end{cases}$

MGF $M_X(t) = E[e^{tX}] = (1-p)\cdot e^{t\cdot 0} + p\cdot e^{t\cdot 1} = 1-p+pe^t$

$$E[X] = \frac{dM_X(t)}{dt}\Big|_{t=0} = pe^t\Big|_{t=0} = p$$

$$E[X^2] = \frac{d^2M_X(t)}{dt^2}\Big|_{t=0} = \frac{d(pe^t)}{dt}\Big|_{t=0} = p$$

$$Var(X) = E[X^2] - (E[X])^2 = p - p^2$$

$$E[X^3] = \frac{d^3M_X(t)}{dt^3}\Big|_{t=0} = p$$

$\therefore$ mean p, variance $(p-p^2)$, $E[X^3] = p$

Problem 4. $\mu = 30$, $\sigma = 15$, $\sigma^2 = 225$,

let $S_{100} = \sum_{i=1}^{100} X_i$ where X_i have above mean and deviation

$$P(S_{100} \le 3500) = P\left(\frac{S_{100} - 100 * \mu}{\sigma\sqrt{100}} \le \frac{3500 - 100 * \mu}{\sigma\sqrt{100}}\right)$$

$$\approx \Phi\left(\frac{3500 - 100\mu}{\sigma\sqrt{100}}\right) = \Phi\left(\frac{10}{3}\right)$$

so $P(S_{100} > 3500) = 1 - \Phi\left(\frac{10}{3}\right)$

Problem 5

(a). $\hat{P}(\text{sell}) = 0.12$

$var(X) = p(1-p) = 0.12 \times 0.88 = 0.1056$

for confidence Interval 0.95, the Problem is to find β for

$$Pr\{\hat{p} - \beta \le p \le p' + \beta\} = 1 - \alpha, \text{ where } \alpha = 0.05$$

$$Pr\left\{\frac{-\beta}{\frac{\sigma}{\sqrt{n}}} \le \frac{p - p'}{\frac{\sigma}{\sqrt{n}}} \le \frac{\beta}{\frac{\sigma}{\sqrt{n}}}\right\} = 1 - \alpha \quad \text{where } \sigma = \sqrt{0.1056}, \; n = 100$$

$$\frac{\beta}{\frac{\sigma}{\sqrt{n}}} = Z_{\alpha/2} \quad \text{where } \alpha/2 = 0.025, \; Z_{\alpha/2} = 1.96$$

$$\beta = 0.06369$$

$\therefore$ confidence level is $[0.0563, 0.1837]$

(b) $\hat{p}(\text{sell}) = 0.18$, sample size $n = 200$,

$Var(X) = p(1-p) = 0.18 \times 0.82 = 0.1476$

$\therefore$ standard error $\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{0.1476}}{\sqrt{200}} = 0.02717$

for 90% confidence interval

$$p \in \left[\hat{p} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \hat{p} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

$$Z_{\alpha/2} = Z_{0.05} = 1.645$$

$$p \in [0.1353, 0.2247]$$

(c) Variance of $\hat{d} = Var(\hat{p}_1) + Var(\hat{p}_2)$

$$= \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}$$

$$= 0.001794$$

std error of $\hat{d} = 0.0424$

(d) 99% confidence level: $(\hat{p}_1 - \hat{p}_2) \pm Z_{0.005} * \sqrt{Var(d)}$, $Z_{0.005} = 2.576$

$[-0.1297, 0.0097]$ $\quad Z_{0.025} = 1.96$

95% CI : $[-0.1431, 0.0231]$ $\quad Z_{0.05} = 1.645$

90% CI : $[-0.1692, 0.0492]$

The difference is unsignificant

Problem 6

$$MSE(\hat{\theta}, \theta) = E[|\hat{\theta} - \theta|^2]$$
$$= E[\hat{\theta}^2] - E[2\hat{\theta}\theta] + E[\theta^2]$$
$$= E[\hat{\theta}^2] - 2\theta E[2\hat{\theta}] + \theta^2$$
$$= E[\hat{\theta}^2] - (E[\hat{\theta}])^2 + (E[\hat{\theta}])^2 - 2\theta E[2\hat{\theta}] + \theta^2$$
$$= (E[\hat{\theta}^2] - (E[\hat{\theta}])^2) + (E[\hat{\theta}] - \theta)^2$$
$$= Var(\hat{\theta}) + (bias(\hat{\theta}, \theta))^2$$

Problem 7.

a). The votes satisfy assumption: 1) 100 sample size $\quad \sigma \; p = 0.25$

2). votes are i.i.d. 3). finite variance (Bernoulli distribution)

$$P(|\hat{p} - p| \geq \delta) = 1 - P(|\hat{p} - p| < \delta) = 1 - P(\hat{p} - \delta < p < \hat{p} + \delta) \approx 0.025$$

$$P(\hat{p} - \delta < p < \hat{p} + \delta) \approxeq 1 - 0.025 \qquad \sigma^2 = p(1-p) = 0.1875$$

$$\delta \approx \frac{\sigma}{\sqrt{n}} Z_{0.0125} = \frac{\sqrt{0.1875}}{\sqrt{100}} \times 2.2414 = 0.0971$$

b) significance level $\alpha = 0.05$

95% CI is $\left[\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$

$= [0.1651, 0.3349]$

Therefore the 95% CI for p contains the true value.