Simon Chen.

HUDM 4125 Due November 18, 2020

HW 7

---

1. Chi square distribution:

a) Use any of the methods shown in class to prove that the mean of a chi-square distribution with n degrees of freedom is equal to n.

b) Derive the variance of a chi-square distribution with n degrees of freedom.

c) Use the CLT to approximate a chi-square distribution with n degrees of freedom with a normal distribution. State explicitly what are the mean and st. dev. of the normal approximation.

d) Use the result from part c) to approximate $P(X > 70)$ where X has chi-square distribution with 50 d.f. Compare your answer with the exact answer obtained with R.

2. In the Smoky Mountains National Park 55 % of the hawks are female. Estimate the number of birds, n, that you would have to catch for there to be a probability of 0.9 of at least 50% female birds in your sample.

3. From Devore's textbook:

   p. 305: Q16

4. From Devore's textbook:

   p. 313: Q37

1. a) mean of χ^2

$f(y) = \frac{1}{(\sqrt{2})^n \sqrt{\frac{n}{2}}} e^{-\frac{y}{2}} y^{\frac{n}{2}-1} \quad y \geq 0$

$\mu'_r = E(y^r) = \int_0^\infty y^r f(y)\,dy = \int_0^\infty \frac{(\frac{1}{2})^{\frac{n}{2}}}{\sqrt{\frac{n}{2}}} y^r e^{-\frac{y}{2}} y^{\frac{n}{2}-1}\,dy$

$= \frac{(\frac{1}{2})^{\frac{n}{2}}}{\sqrt{\frac{n}{2}}} \int_0^\infty e^{-\frac{y}{2}} y^{\frac{n}{2}+r-1}\,dy$

$= \frac{(\frac{1}{2})^{\frac{n}{2}}}{\sqrt{\frac{n}{2}}} \frac{\sqrt{\frac{n}{2}+r}}{(\frac{1}{2})^{\frac{n}{2}+r}}$

$\mu'_r = \frac{\sqrt{\frac{n}{2}+r}}{\sqrt{\frac{n}{2}}} 2^r = 2^r (\frac{n}{2}+r-1)(\frac{n}{2}+r-2) - (\frac{n}{2})$

$\mu'_r = n(n+2)\ldots(n+2r-2)$

so $\mu'_r = (n+r-2) = n \quad \therefore$ mean $= n$.

b) $\mu'_2 = n(n+2) \quad Var(y) = \mu'_2 - (\mu'_1)^2$

$= n(n+2) - n^2$

$= n^2 + 2n - n^2 = 2n$

c) by CLT

Let $x_1\, x_2 \ldots x_n$ be i.i.d. random varibles with

$E(x_i) = \mu < \infty$ & $var(x) = \sigma^2 < \infty$

then $Z_n = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$\therefore \sum_{i=1}^{n} Z_n^2 \sim \chi^2_n$ d.b.

For example: the sum of n square of std Normal varibles follows χ^2 with n degrees of freedom.

so Z_n has $N(0,1)$ with mean $\mu_1 = 0$ $\quad var(\sigma^2) = 2$

d) $p(X>70)$ where $X \sim \chi^2_{50df}$.
by using chi-square table. $p(X>70)=0.0323741$

2. CLT to estimate n.

$$p\left(\frac{S_n/n-\mu}{\sigma/\sqrt{n}} \le z\right)=0.9 \Rightarrow \frac{0.5-0.55}{0.497/\sqrt{n}}=1.2816$$

$$\Rightarrow n=\left(\frac{1.2816\times 0.497}{-0.05}\right)^2 \approx 163$$

day 1

3. $p(\bar{x}_1 \le 11)= p\left(Z\le \frac{11-10}{2/\sqrt{5}}\right)=p(Z\le 1.12)=0.8686.$

$p(\bar{x}_2 \le 11)= p\left(Z\le \frac{11-10}{2/\sqrt{6}}\right)=p(Z\le 1.22)=0.8888$

$p=p(\bar{x}_1\le 11)\times p(\bar{x}_2\le 11)=0.8686\times 0.8888=0.7720$

4. X_1 X_2 for speed of 2 plants.
Y for distance between them after 2 hours

$Y=2X_1-2X_2+10$

$X_1\sim N(520,(10)^2)$

$X_2\sim N(500,(10)^2)$

$E(Y)=E[2X_1-2X_2+10]=2\mu_1-2\mu_2+10=2\times520-2\times500+10$
$=50$

$SD(Y)=\sqrt{V(Y)}$
$=\sqrt{V(2X_1-2X_2+10)}$
$=\sqrt{4\sigma_1^2+4\sigma_2^2}=\sqrt{4(10^2)+4(10)^2}=28.28427$

a) $p(Y>0)=p\left(\frac{Y-E(Y)}{SD(Y)}>\frac{0-50}{28.28427}\right)=p(Z>-1.77)=1-\Phi(-1.77)$ $=0.9616$

b) $p(|Y|\le 10)=p(-10\le Y\le 10)=p\left(\frac{-10-50}{28.28427}\le\frac{Y-E(Y)}{SD(Y)}\le\frac{10-50}{28.28427}\right)$
$=\Phi(-1.41)-\Phi(-2.12)=0.0623$