Statistical Inference Chapter 5

Gallant Tsao

August 13, 2024

1. Let X be the number of color blinded people in that population of size n. Then $X \sim \text{Binomial}(n, 0.01)$. We want

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \binom{n}{0}(0.01)^0(0.99)^n \\ &= 1 - 0.99^n \\ &> 0.95. \end{aligned}$$

   For the inequality above to be true, we need $n > \log_{0.99}(0.05) \implies n \geq 299$.

2. (a)

3. From the definition, we can see that $Y_i \sim \text{Bernoulli}(1 - F(\mu))$. Since $\sum_{i=1}^n Y_i$ is a sum of independent Bernoulli random variables, we can get that

$$\sum_{i=1}^n Y_i \sim \text{Binomial}(n, 1 - F(\mu)).$$

4. (a)

5. Let $Y = \sum_{i=1}^n X_i$. We can see that $\bar{X} = (1/n)Y$ is a scale transformation. Then the pdf of $\bar{X}$ is

$$f_{\bar{X}}(x) = \frac{1}{1/n} f_Y(\frac{x}{1/n}) = n f_X(nx).$$

6. (a) Set $Z = X - Y$, $W = X$. Then the Jacobian of (X, Y) to (Z, W) is

7. (a)

8. 5.8

9. 5.9

10. 5.10

11. 5.11

12. First note that since $X_1, \cdots, X_n$ are $N(0, 1)$, $\bar{X} \sim N(0, \frac{1}{n})$. In particular, $Y_1 = |\bar{x}|$ is a folded normal distribution with mean 0 and variance $\frac{1}{n}$. Then we can directly obtain that the expectation for Y_1 is

$$\mathbb{E}[Y_1] = \sqrt{\frac{2}{\pi n}}.$$

   Similarly, we can see the Y_2 is the average of n folded normals with mean 0 and variance 1, so

$$\mathbb{E}[Y_2] = \frac{1}{n}(n\sqrt{\frac{2}{\pi}}) = \sqrt{\frac{2}{\pi}}.$$

   From the above we can see clearly that $\mathbb{E}[Y_1] \leq \mathbb{E}[Y_2]$.

13. 5.13

14. 5.14

15. 5.15

16. (a)
$$(X_1-1)^2+\big(\frac{X_2-2}{2}\big)^2+\big(\frac{X_3-3}{3}\big)^2\sim\chi^2_3.$$

(b)
$$\frac{X_1-1}{\sqrt{\big(\frac{X_2-2}{2}\big)^2+\big(\frac{X_3-3}{3}\big)^2/2}}\sim t_2.$$

(c) Since $F_{1,2}\sim T_2^2$, squaring the random variable from part (b) gives the result.

17. (a)
$$f_X(x)=\frac{\Gamma(\frac{p+q}{2})p^{p/2}q^{q/2}x^{p/2-1}}{\Gamma(\frac{p}{2})\Gamma(\frac{q}{2})(q+px)^{(p+q)/2}},\ x>0.$$

(b) We can write $X=\frac{U/p}{V/q}, U\sim\chi^2_p, V\sim\chi^2_q$. Firstly, note that for $Y\sim\chi^2_n$,

$$\begin{aligned}\mathbb{E}[Y^k]&=\int_0^\infty x^k\cdot\frac{x^{n/2-1}e^{-x/2}}{2^{n/2}\Gamma(\frac{n}{2})}\,dx\\&=\int_0^\infty\frac{x^{n/2+k-1}e^{-x/2}}{2^{n/2}\Gamma(\frac{n}{2})}\,dx\\&=\frac{\Gamma(\frac{n}{2}+k)2^k}{\Gamma(\frac{n}{2})}\int_0^\infty\frac{x^{n/2+k-1}e^{-x/2}}{2^{n/2+k}\Gamma(\frac{n}{2}+k)}\,dx\\&=\frac{\Gamma(\frac{n}{2}+k)2^k}{\Gamma(\frac{n}{2})},\ k>-\frac{n}{2}.\end{aligned}$$

Plugging in $k=-1$ into V gives $\mathbb{E}[V^{-1}]=\frac{1}{q-2}$. Then

$$\begin{aligned}E[X]&=\mathbb{E}\big[\frac{U/p}{V/q}\big]\\&=\frac{1}{pq}\cdot\mathbb{E}[U]\mathbb{E}[V]\\&=\frac{q}{q-2},\ q>2.\end{aligned}$$

As for the variance, first note that

$$\mathbb{E}[X^2]=\frac{q^2}{p^2}\mathbb{E}[U^2]\mathbb{E}[V^{-2}].$$

From the equation above, plugging in $k=2$ and $k=-2$ repectively gives

$$\mathbb{E}[U^2]=\frac{4\Gamma(\frac{p}{2}+2)}{\Gamma(\frac{p}{2}+2)}=4\big(\frac{p}{2}+1\big)\frac{p}{2}=2p+p^2,$$

$$\mathbb{E}[V^{-2}]=\frac{2^{-2}\Gamma(\frac{q}{2}-2)}{\Gamma(\frac{q}{2})}=\frac{1}{4(\frac{q}{2}-1)(\frac{q}{2}-2)}=\frac{1}{(q-2)(q-4)},\ q>4.$$

Finally we get that

$$\begin{aligned}\operatorname{Var} X &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= \frac{q^2(2p+p^2)}{p^2(q-2)(q-4)} - \frac{q^2}{(q-2)^2} \\ &= \frac{2q^2(p+q-2)}{p(q-2)^2(q-4)},\ q>4.\end{aligned}$$

(c) Let U and V be as given above. Then

$$\frac{1}{X} = \frac{V/q}{U/p} \sim F_{q,p}.$$

(d)

18. First note that if $X \sim t_p$, $X \sim \frac{Z}{\sqrt{V/p}}$, $V \sim \chi_p^2$, with Z and V independent. The moments for V can be obtained from the equation in part (b) of Exercise 5.17.

(a)

$$\mathbb{E}[X] = \sqrt{p}\mathbb{E}[Z]\mathbb{E}[\frac{1}{\sqrt{V}}] = 0.$$

$$\operatorname{Var} X = \mathbb{E}[X^2] = \mathbb{E}[\frac{Z^2}{V/p}] = p\mathbb{E}[Z^2]\mathbb{E}[V^{-1}] = \frac{p}{p-2},\ p>2.$$

(b)

$$X^2 \sim \frac{Z^2}{V/p} \sim \frac{\chi_1^2/1}{\chi_p^2/p} \sim F_{1,p}.$$

(c)

19. (a) $\chi_p^2 \sim \chi_q^2 + \chi_d^2$ where χ_q^2, χ_d^2 are independent random variables and $d = p-q$. Since χ_d^2 is a strictly positive random variable, for all $a>0$,

$$P(\chi_p^2 > a) = P(\chi_q^2 + \chi_d^2 > a) > P(\chi_q^2 > a).$$

(b)

20. 5.20

21. 5.21