Stat 588 HW6

① $X_i \overset{iid}{\sim} \text{Bernoulli}(\theta) \quad (i = 1, 2, \dots, n)$

$f(x;\theta) = \theta^x (1-\theta)^{1-x}$ for $x = 0, 1$.

a) $\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{\mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]}{n}$

$= \frac{n\,\mathbb{E}[X_i]}{n} = \frac{n\theta}{n} = \theta \quad (i = 1, 2, \dots, n)$

$(\mathbb{E}[X_i] = \theta)$

b) $\text{Var}(\bar{X}) = \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right)$

$\text{Var}\left(\sum_{i=1}^n X_i\right) = \mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] - \left(\mathbb{E}\left[\sum_{i=1}^n X_i\right]\right)^2.$

$\mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n X_i X_j\right]$

$= \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i X_j].$

$\left(\mathbb{E}\left[\sum_{i=1}^n X_i\right]\right)^2 = \left(\sum_{i=1}^n \mathbb{E}[X_i]\right)^2 = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i]\mathbb{E}[X_j]$

$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i X_j] - \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i]\mathbb{E}[X_j]$

$= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) = \sum_{i=1}^n \text{Cov}(X_i, X_i) = \sum_{i=1}^n \text{Var}(X_i)$

We have

$$\mathbb{E}[X_i] = \sum_{x=0}^{1} \theta^x (1-\theta)^{1-x} x = \theta$$

$$\mathbb{E}[X_i^2] = \sum_{x=0}^{1} x^2 \theta^x (1-\theta)^{1-x} = \theta$$

$$\Rightarrow Var(X_i) = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = \theta - \theta^2 = \theta(1-\theta)$$

Hence,

$$Var(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} n \theta (1-\theta) = \frac{\theta(1-\theta)}{n}$$

c) Theorem: If $X_1, \dots, X_n$ are RVs and $Y = \sum_{i=1}^{n} a_i X_i$, where a_i are constant, then: $\mathbb{E}[Y] = \sum_{i=1}^{n} a_i \mathbb{E}[X_i]$ and $Var(Y) = \sum_{i=1}^{n} a_i^2 var(X_i) + 2 \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j)$.

(2) $X_i \sim Bernoulli(\theta_1)$; $Y_j \sim Bernoulli(\theta_2)$

$(i = 1, \dots, n_1 \; ; \; j = 1, \dots, n_2)$

a) $\mathbb{E}[\bar{X} - \bar{Y}] = \mathbb{E}[\bar{X}] - \mathbb{E}[\bar{Y}] = \theta_1 - \theta_2$

b) $Var(\bar{X} - \bar{Y}) = \mathbb{E}[(\bar{X} - \bar{Y})^2] - (\mathbb{E}[\bar{X} - \bar{Y}])^2$

$$\mathbb{E}[(\bar{X} - \bar{Y})^2] = \mathbb{E}[\bar{X}^2 - 2\bar{X}\bar{Y} + \bar{Y}^2]$$

$$= \mathbb{E}[\bar{X}^2] - 2\mathbb{E}[\bar{X}\bar{Y}] + \mathbb{E}[\bar{Y}^2]$$

$$(\mathbb{E}[\bar{X} - \bar{Y}])^2 = (\mathbb{E}[\bar{X}] - \mathbb{E}[\bar{Y}])^2$$

$$= (\mathbb{E}[\bar{X}])^2 - 2\mathbb{E}[\bar{X}]\mathbb{E}[\bar{Y}] + (\mathbb{E}[\bar{Y}])^2$$

$$\Rightarrow Var(\bar{X} - \bar{Y}) = \mathbb{E}[\bar{X}^2] - (\mathbb{E}[\bar{X}])^2 + \mathbb{E}[\bar{Y}^2] - (\mathbb{E}[\bar{Y}])^2 - 2(\mathbb{E}[\bar{X}\bar{Y}] - \mathbb{E}[\bar{X}]\mathbb{E}[\bar{Y}])$$

(because X and Y are independent)

$$= Var(\bar{X}) + Var(\bar{Y}) + 0$$

$$= \frac{\theta_1(1-\theta_1)}{n} + \frac{\theta_2(1-\theta_2)}{n}$$

(3) Let $X_1, \dots, X_n$ be a random sample from an infinite population with mean μ and variance σ^2.
Let $\bar{X} = \frac{X_1 + \dots + X_n}{n}$ be the sample mean.

Then $Var(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) = \frac{Var(X_i)}{n}$

a) When $n = 30$ increases to $n' = 120$, then $n' = 4n$.
Moreover, $\sqrt{Var(\bar{X})}$ will decrease by 2 times.
The factor $\sqrt{Var(\bar{X})}$ is divided by is 2.

b) When $n = 450$ decreases to $n' = 50$, then $n' = \frac{1}{9} n$.
Moreover, $\sqrt{Var(\bar{X})}$ will increase by 3 times
The factor $\sqrt{Var(\bar{X})}$ is multiplied by is 3.

(4) $n = 100$, $\mu = 75$, $\sigma^2 = 256$.

a) • By law of Large Numbers:
For any positive constant c, the probability that $\bar{X}$ will take on a value between $\mu - c$ and $\mu + c$ is at least $1 - \frac{\sigma^2}{nc^2}$. When $n \to \infty$, this probability approaches 1.

• In our case, $c = 8$. Then
$\mathbb{P}(67 \leq \bar{X} \leq 83) \geq 1 - \frac{256}{100 \cdot 8^2} = 0.96$.

b) $\mathbb{P}(67 \leq \bar{X} \leq 83) = \mathbb{P}\left(\frac{67 - \mu}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{83 - \mu}{\sigma/\sqrt{n}}\right)$

$= \mathbb{P}\left(-5 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 5\right) = \mathbb{P}\left(\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right| \leq 5\right)$

$= \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 5\right) - \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq -5\right)$

$= 0.9999997 - (1 - 0.999997) = 0.9999994$

(5) Theorem:

If S^2 is the variance of a sample of size n from a normal distribution with mean μ and variance σ^2, then

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}.$$

Let X be a random variable that follows $\chi^2_{(n-1)}$.

Then

$$f(x) = \begin{cases} \frac{1}{2^{v/2}\Gamma(v/2)} x^{\frac{v-2}{2}} e^{-x/2} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$M_X(t) = (1-2t)^{-v/2}$$

$$\mathbb{E}[X] = \frac{\partial}{\partial t} M_X(t)\Big|_{t=0} = v(1-2t)^{\frac{-v-2}{2}}\Big|_{t=0} = v$$

$$\mathbb{E}[X^2] = \frac{\partial^2}{\partial t^2} M_X(t)\Big|_{t=0} = \frac{\partial}{\partial t}\left(v(1-2t)^{\frac{-v-2}{2}}\right)\Big|_{t=0}$$

$$= v(v+1)(1-2t)^{\frac{-v}{2}-2}\Big|_{t=0} = v(v+1).$$

Then:

$$\mathbb{E}\left[\frac{(n-1)S^2}{\sigma^2}\right] = v$$

$$Var\left[\frac{(n-1)S^2}{\sigma^2}\right] = v(v+\ ) - v^2 = 2v$$

a) $\mathbb{E}\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{(n-1)\mathbb{E}[S^2]}{\sigma^2} = v = n-1.$

$\Rightarrow \mathbb{E}[S^2] = \sigma^2$

b) $Var\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{(n-1)^2 Var(S^2)}{\sigma^4} = 2v = 2(n-1)$

$\Rightarrow Var(S^2) = \frac{2(n-1)\sigma^4}{(n-1)^2} = \frac{2\sigma^4}{(n-1)}$

(6) As F is a RV following F distribution, we can write F as:

$$F = \frac{U/v_1}{V/v_2},$$

where U and V are 2 RVs, independent, and following Chi-square distribution with v_1 d.f and v_2 d.f, respectively.

Then, $\mathbb{E}(F) = \mathbb{E}\left(\frac{U/v_1}{V/v_2}\right) \overset{ind.}{=} \mathbb{E}\left(\frac{U}{v_1}\right) \cdot \mathbb{E}\left(\frac{v_2}{V}\right)$

$\mathbb{E}\left(\frac{U}{v_1}\right) = \frac{1}{v_1}\mathbb{E}(U) = \frac{v_1}{v_1} = 1.$

$\mathbb{E}\left(\frac{v_2}{V}\right) = v_2\,\mathbb{E}\left[\frac{1}{V}\right] = \frac{v_2}{v_2-2} \quad (v_2 > 2)$

Hence, $E(F) = 1 \cdot \frac{v_2}{v_2-2} = \frac{v_2}{v_2-2} \quad (v_2 > 2).$

⑦ $T \sim$ t-distribution.

$$f(t) = \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{\pi\upsilon}\,\Gamma\left(\frac{\upsilon}{2}\right)}\left(1+\frac{t^2}{\upsilon}\right)^{-\frac{(\upsilon+1)}{2}} \quad (-\infty < t < \infty)$$

$X = g(T) = T^2$ is a transformation with $T = g^{-1}(X) = \sqrt{X}$ and $\frac{dT}{dX} = \frac{1}{2\sqrt{X}}$. Then

$$f_X(x) = f_T(g^{-1}(x))\left|\frac{dT}{dX}\right|$$

$$= \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\upsilon}{2}\right)} \cdot \frac{1}{\sqrt{\upsilon}}\left(1+\frac{x}{\upsilon}\right)^{-\frac{(\upsilon+1)}{2}} \cdot \frac{1}{2x^{1/2}}$$

$$= \frac{1}{2}\,\frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{\upsilon}{2}\right)} \cdot \left(\frac{1}{\upsilon}\right)^{\frac{1}{2}}\left(1+\frac{x}{\upsilon}\right)^{-\frac{1}{2}(1+\upsilon)} x^{\frac{1}{2}-1},$$

which is of the same form as the density function of a F-distribution with nominator d.f $v_1 = 1$ and denominator $v_2 = v$.

Note that

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} t^{-1/2} e^{-t}\, dt \xrightarrow{u^2 := t} \int_0^{+\infty} (u^2)^{-1/2} e^{-u^2}\, 2u\,du$$

$$= 2\int_0^{+\infty} e^{-u^2}\, du = \sqrt{\pi}.$$

Hence $X \sim F(1, v)$.

(8) Random sample $X \sim \exp(x, \theta)$. Then $\mathbb{E}[X] = \theta$; $Var(X) = \theta^2$.

Hence, $\mathbb{E}[\bar{X}] = \theta$ and $Var(\bar{X}) = \frac{\theta^2}{n}$.

By Central Limit Theorem, $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1)$.

Now we have:

$$\mathbb{P}(\bar{X} \geq 4.5) = \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{4.5 - \mu}{\sigma/\sqrt{n}}\right)$$

$$= \mathbb{P}\left(Z \geq \frac{4.5 - 4}{4/\sqrt{225}}\right) = \mathbb{P}(Z \geq 1.875)$$

(9)

$$\mathbb{P}(|X - \mu| \geq 3) = \mathbb{P}(X - \mu \geq 3 \text{ or } X - \mu \leq -3)$$

$$= \mathbb{P}(X - \mu \geq 3) + \mathbb{P}(X - \mu \leq -3)$$

$$= \mathbb{P}\left(\frac{X - \mu}{\sigma/\sqrt{n}} \geq \frac{3}{\sigma/\sqrt{n}}\right) + \mathbb{P}\left(X - \mu \leq \frac{-3}{\sigma/\sqrt{n}}\right)$$

$$= \mathbb{P}\left(\frac{X - \mu}{0.5} \geq 6\right) + \mathbb{P}\left(\frac{X - \mu}{0.5} \leq -6\right)$$

$$= \mathbb{P}(Z \geq 6) + \mathbb{P}(Z \leq -6)$$

with $Z \sim \mathcal{N}(0,1)$

(10) $\mathbb{P}(\text{Rejecting } \sigma^2 = 25) = \mathbb{P}(S^2 \geq 54.668 \text{ or } S^2 \leq 12.102)$

$= \mathbb{P}(S^2 \geq 54.668) + \mathbb{P}(S^2 \leq 12.102)$

$= \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} \geq \frac{(n-1).(54.668)}{\sigma^2}\right) + \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} \leq \frac{(n-1)(12.102)}{\sigma^2}\right)$

$= \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} \geq 32.8008\right) + \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} \leq 7.2612\right)$

with $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$