HW6 solutions

1. (a) Since $P(X_i \le y) = \frac{y}{H}$ for $y = 1, \dots, H$,

$$P(\max_{1 \le i \le n} X_i \le y) = P(X_1 \le y, X_2 \le y, \dots, X_n \le y)$$
$$= \prod_{i=1}^{n} P(X_i \le y) = \left(\frac{y}{H}\right)^n.$$

Then we have
$$P(\max_{1 \le i \le n} X_i = y) = P(\max_{1 \le i \le n} X_i \le y) - P(\max_{1 \le i \le n} X_i \le y-1)$$
$$= \left(\frac{y}{H}\right)^n - \left(\frac{y-1}{H}\right)^n$$
$$= \frac{y^n - (y-1)^n}{H^n} \quad \text{for } y = 1, \dots, H.$$

(b) Suppose $E[g(Y)] = 0$ for all $H \in \mathbb{Z}^+$ where $Y = \max_{1 \le i \le n} X_i$.

$$\Downarrow$$

$$\sum_{y=1}^{H} g(y) \frac{y^n - (y-1)^n}{H^n} = 0 \quad \text{for } H = 1, 2, 3, \dots \qquad (*)$$

Then we show that $g(y) = 0$ for $y = 1, 2, 3, \dots$ by induction.

<u>Basis step</u>: Using $(*)$ with $H = 1$, $g(1) \cdot \frac{1}{1} = 0 \Rightarrow g(1) = 0$

<u>Inductive step</u>: Suppose $g(1) = \dots = g(k) = 0$ and show $g(k+1) = 0$.

Using $(*)$ with $H = k+1$, we have

$$\underbrace{g(1)}_{=0} \frac{1}{(k+1)^n} + \underbrace{g(2)}_{=0} \frac{2^n - 1}{(k+1)^n} + \dots + \underbrace{g(k)}_{=0} \frac{k^n - (k-1)^n}{(k+1)^n} + g(k+1) \frac{(k+1)^n - k^n}{(k+1)^n} = 0$$

$$\Rightarrow g(k+1) \frac{(k+1)^n - k^n}{(k+1)^n} = 0 \Rightarrow g(k+1) = 0.$$

Thus, $g(y) = 0$ for $y = 1, 2, 3, \dots$

This shows that the family of pdfs of $\max_{1 \le i \le n} X_i$ is complete

since $\{\omega : g(X(\omega)) = 0\} \supset \{\omega : X(\omega) \in \mathbb{Z}^+\}$ so that

$$P(g(X) = 0) = P(X \in \mathbb{Z}^+) = 1.$$

1. (EC - part 1) The joint pmf of $X_1,\dots,X_n$ is

$$f(x_1,\dots,x_n) = \prod_{i=1}^{n} I_{\{1,\dots,H\}}(x_i) = I_{\{1,\dots,H\}}\left(\max_{1\le i\le n} x_i\right)$$

So $\max_{1\le i\le n} X_i$ is sufficient for H by the Factorization Theorem.

There is an unbiased estimator of H since

$$E[2X_1-1] = 2E[X_1]-1 = 2\frac{(1+H)}{2} - 1 = H$$

So, since $\max_{1\le i\le n} X_i$ is complete and sufficient for H,

Theorem L11.5 implies that $\phi\left(\max_{1\le i\le n} X_i\right) = E\left[2X_1-1 \,\middle|\, \max_{1\le i\le n} X_i\right]$

is the UMVUE of H (and consequently a UMVUE exists).

(EC - part 2) Let $U(Y)$ be the UMVUE of H.

$$E[U(Y)] = H \Leftrightarrow \sum_{y=1}^{H} U(y)\,\frac{y^n-(y-1)^n}{H^n} = H$$

$$\Rightarrow \sum_{y=1}^{H} U(y)\,\frac{y^n-(y-1)^n}{H^{n+1}} = 1$$

$$\Rightarrow \sum_{y=1}^{H} U(y)\,\frac{y^n-(y-1)^n}{y^{n+1}-(y-1)^{n+1}}\cdot\frac{y^{n+1}-(y-1)^{n+1}}{H^{n+1}} = 1$$

$$\Rightarrow \sum_{y=1}^{H} U(y)\,\frac{y^{m-1}-(y-1)^{m-1}}{y^m-(y-1)^m}\,\frac{y^m-(y-1)^m}{H^m} = 1 \quad \text{where } m=n+1$$

$$\Rightarrow E\left[U(Y)\frac{Y^{m-1}-(Y-1)^{m-1}}{Y^m-(Y-1)^m}\right] = 1$$

$$\Rightarrow E\left[U(Y)\frac{Y^{m-1}-(Y-1)^{m-1}}{Y^m-(Y-1)^m} - 1\right] = 0$$

$$\Rightarrow U(Y)\frac{Y^{m-1}-(Y-1)^{m-1}}{Y^m-(Y-1)^m} - 1 = 0 \quad \text{with probability 1 by completeness}$$

$$\Rightarrow U(Y)\frac{Y^{m-1}-(Y-1)^{m-1}}{Y^m-(Y-1)^m} = 1$$

$$\Rightarrow U(Y) = \frac{Y^m-(Y-1)^m}{Y^{m-1}-(Y-1)^{m-1}} = \frac{Y^{n+1}-(Y-1)^{n+1}}{Y^n-(Y-1)^n}$$

is the UMVUE for H where $Y = \max_{1\le i\le n} X_i$.

Here is the direct calculation of the Rao-Blackwell estimator.

$$P(X_1 = x \mid X_{(n)} = y) = \frac{P(X_1 = x \text{ and } X_{(n)} = y)}{P(X_{(n)} = y)}$$

$$P(X_1 = x \text{ and } X_{(n)} = y) = \begin{cases} P(X_1 = x, \max_{2 \le i \le n} X_i = y) & \text{if } x < y \\ P(X_1 = y, X_2 \le y, \ldots, X_n \le y) & \text{if } x = y \end{cases}$$

$$= \begin{cases} P(X_1 = x)\, P(\max_{2 \le i \le n} X_i = y) & \text{for } x = 1, \ldots, y-1 \\ P(X_1 = y)\, P(X_2 \le y) \cdots P(X_n \le y) & \text{for } x = y \end{cases}$$

$$= \begin{cases} \frac{1}{H} \cdot \frac{y^{n-1} - (y-1)^{n-1}}{H^{n-1}} & \text{for } x = 1, \ldots, y-1 \\ \frac{1}{H} \cdot \frac{y^{n-1}}{H^{n-1}} & \text{for } x = y \end{cases}$$

$$= \begin{cases} \frac{y^{n-1} - (y-1)^{n-1}}{H^n} & \text{for } x = 1, \ldots, y-1 \\ \frac{y^{n-1}}{H^n} & \text{for } x = y \end{cases}$$

Since $P(X_{(n)} = y) = \frac{y^n - (y-1)^n}{H^n}$, we have

$$P(X_1 = x \mid X_{(n)} = y) = \begin{cases} \frac{y^{n-1} - (y-1)^{n-1}}{y^n - (y-1)^n} & \text{for } x = 1, \ldots, y-1 \\ \frac{y^{n-1}}{y^n - (y-1)^n} & \text{for } x = y \end{cases}$$

Then $$E[X_1 \mid X_{(n)} = y] = \sum_{x=1}^{y-1} x \, \frac{y^{n-1} - (y-1)^{n-1}}{y^n - (y-1)^n} + y \cdot \frac{y^{n-1}}{y^n - (y-1)^n}$$

$$= \frac{(y-1)y}{2} \cdot \left(\frac{y^{n-1} - (y-1)^{n-1}}{y^n - (y-1)^n} \right) + \frac{y^n}{y^n - (y-1)^n}$$

$$= \frac{1}{2(y^n - (y-1)^n)} \cdot \left\{ y(y-1)(y^{n-1} - (y-1)^{n-1}) + 2y^n \right\}$$

$$= \frac{1}{2(y^n - (y-1)^n)} \cdot \left\{ y^n(y-1) - y(y-1)^n + 2y^n \right\}$$

$$= \frac{1}{2(y^n-(y-1)^n)} \left\{ y^{n+1} - \cancel{y^n} - (y-1+1)(y-1)^n + \cancel{2}y^n \right\}$$

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

$$= \frac{y^{n+1}-(y-1)^{n+1}}{2(y^n-(y-1)^n)} + \frac{1}{2} .$$

So $$E[2X_1 - 1 \mid X_{(n)} = y] = 2E[X_1 \mid X_{(n)} = y] - 1$$

$$= \frac{y^{n+1}-(y-1)^{n+1}}{y^n-(y-1)^n}$$

and by Theorem L11.5,

$$E[2X_1 - 1 \mid X_{(n)}] = \frac{X_{(n)}^{n+1} - (X_{(n)}-1)^{n+1}}{X_{(n)}^n - (X_{(n)}-1)^n}$$

is the UMVUE of H.

2. (a) Let x be the number of red marbles selected in the sample. Then H_0 is rejected if $x=0$ so $\{0\}$ is the rejection region.

(b) $X \sim \text{Binomial}\left(n=4, p=\frac{M}{10}\right)$

$$P(\text{Type I error}) = P_{M=6}(X=0) = \binom{4}{0}.6^0 .4^{4-0} = .4^4 = .0256$$

(c) $$P(\text{Type II error with } M=5) = 1-P_{M=5}(X=0) = 1-\binom{4}{0}.5^0 .5^{4-0} = 1-.5^4 = 1-.0625 = .9375$$

(d)

$$\beta(M) = \begin{cases} 1 & \text{if } M=0 \\ (.9)^4 = .6561 & \text{if } M=1 \\ (.8)^4 = .4096 & \text{if } M=2 \\ (.7)^4 = .2401 & \text{if } M=3 \\ (.6)^4 = .1296 & \text{if } M=4 \\ (.5)^4 = .0625 & \text{if } M=5 \\ (.4)^4 = .0256 & \text{if } M=4 \\ (.3)^4 = .0081 & \text{if } M=3 \\ (.2)^4 = .0016 & \text{if } M=2 \\ (.1)^4 = .0001 & \text{if } M=1 \\ 0^4 = 0 & \text{if } M=0 \end{cases}$$

(e) $\sup_{M \in [6,10]} \beta(M) = .0256$ is the size of the test

(f) Since the size .0256 is less than or equal to .05, this is a level $\alpha = .05$ test.

3. (a) The likelihood function

$$L(\mu;\underline{x}) = (2\pi)^{-n/2} e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\mu)^2}$$

is increasing when $\mu < \bar{x}$ and decreasing when $\mu > \bar{x}$ since

$$\ell(\mu;\underline{x}) = \ln L(\mu;\underline{x}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\Sigma(x_i-\mu)^2$$

also has this shape because $\frac{d\ell}{d\mu} = \frac{\cancel{2}}{\cancel{2}}\Sigma(x_i-\mu) = (\Sigma x_i - n\mu)$

$$= n(\bar{x}-\mu)$$

is positive when $\mu < \bar{x}$ and negative when $\mu > \bar{x}$.

So $\sup_{\mu \le 0} L(\mu;\underline{x}) = L(0;\underline{x}) = (2\pi)^{-n/2} e^{-\frac{1}{2}\Sigma x_i^2}$

and $\sup_{\mu \in [0,\infty)} L(\mu;\underline{x}) = \begin{cases} L(0;\underline{x}) & \text{if } \bar{x} < 0 \\ L(\bar{x};\underline{x}) & \text{if } \bar{x} \ge 0 \end{cases}$

and the likelihood ratio is $\lambda(\underline{x}) = \begin{cases} 1 & \text{if } \bar{x} < 0 \\ \underbrace{\dfrac{e^{-\frac{1}{2}\Sigma x_i^2}}{e^{-\frac{1}{2}\Sigma(x_i-\bar{x})^2}}}_{\shortparallel} & \text{if } \bar{x} \ge 0 \end{cases}$

$$e^{-\frac{1}{2}(\Sigma x_i^2 - \Sigma(x_i-\bar{x})^2)}$$

$$= e^{-\frac{1}{2}(\Sigma x_i^2 - \Sigma x_i^2 + n\bar{x}^2)}$$

$$= e^{-\frac{n\bar{x}^2}{2}} = e^{-\frac{(\Sigma x_i)^2}{2n}}$$

Note that $\lambda(\underline{x})$ ~~which~~ is a decreasing function of Σx_i.

We reject H_0 if $\lambda(\underline{x}) \le c \Leftrightarrow \Sigma x_i \ge K$.

(b) Since $X_i \overset{\text{indep.}}{\sim} N(\mu,1)$, $\Sigma X_i \sim N(n\mu, n) \Rightarrow \frac{\Sigma X_i - n\mu}{\sqrt{n}} \sim N(0,1)$.

If H_0 is true, $\mu = 0$ and we find K such that

$$P(\Sigma X_i \ge K) = .05.$$

$$\Updownarrow$$

$$P\Big(\underbrace{\frac{\Sigma X_i}{\sqrt{n}}}_{\sim N(0,1)} \ge \frac{K}{\sqrt{n}}\Big) = .05$$

$$\Updownarrow$$

$$\frac{K}{\sqrt{n}} = \Phi^{-1}(.95) \Leftrightarrow K = \sqrt{n}\,\Phi^{-1}(.95) \approx \boxed{1.645\sqrt{n}}.$$

(c) $$\beta(\mu) = P\left(\sum X_i \geq \sqrt{n}\, \Phi^{-1}(.95) \right)$$
$$= P\left(\frac{\sum X_i - n\mu}{\sqrt{n}} \geq \frac{\sqrt{n}\,\Phi^{-1}(.95) - n\mu}{\sqrt{n}} \right)$$
$$= P\left(\underbrace{\frac{\sum X_i - n\mu}{\sqrt{n}}}_{\sim N(0,1)} \geq \Phi^{-1}(.95) - \sqrt{n}\mu \right)$$
$$= 1 - \Phi\left(\Phi^{-1}(.95) - \sqrt{n}\mu \right) \approx \boxed{1 - \Phi(1.645 - \sqrt{n}\mu)}.$$

(d) Since $\beta(\mu)$ is increasing in μ, we need to find the smallest n such that the power of the test is at least .9 when $\mu = 1$. Then that n will work for all $\mu > 1$ since $\beta(\mu) \geq \beta(1) \geq .9$.

So, we have
$$1 - \Phi(1.645 - \sqrt{n} \cdot 1) \geq .90$$
$$\Updownarrow$$
$$\Phi(1.645 - \sqrt{n}) \leq .10$$
$$\Updownarrow$$
$$1.645 - \sqrt{n} \leq \Phi(.10) \approx -1.282$$
$$\Updownarrow$$
$$2.927 = 1.645 + 1.282 \leq \sqrt{n}$$
$$\Updownarrow$$
$$8.567 \leq n.$$

Thus, the smallest n which satisfies the condition is $\boxed{n = 9}$.