DAY 1, QU 1.

(HXZ-1) $X_i \overset{iid}{\sim} N(\mu, 1)$, $i=1,\ldots,n$ — unobserved

$Y_i = I(X_i < 0)$, $i=1,\ldots,n$ — observed

(a) Note that $Y_i \overset{iid}{\sim} \text{Bernoulli}(p)$, $i=1,\ldots,n$,
where $p = P(X<0) = \Phi(-\mu)$, and $\Phi(\cdot)$ is the standard normal cdf.
Because the MLE for p is $\bar{Y}$. By the invariance property of MLE, the MLE for μ is $-\Phi^{-1}(\bar{Y})$, where $\Phi^{-1}(\cdot)$ is the inverse function of $\Phi(\cdot)$.

(b) $\sum_{i=1}^{n} Y_i$ is a sufficient statistic for p. Let $T(\underset{\sim}{Y}) = \sum_{i=1}^{n} Y_i$.
It follows that $f_{\underset{\sim}{Y}|T}(\underset{\sim}{Y}|T)$ is free of p, and thus is also free of $\mu = -\Phi^{-1}(p)$. Therefore, T is a sufficient statistic for μ.

Because $T \sim \text{Binomial}(n, p)$, and the distribution family $\{\text{Binomial}(n,p),\ p\in[0,1]\}$ is complete, T is a complete statistic.

(c) Test $H_0: \mu \le \mu_0$ versus $H_1: \mu > \mu_0$.
For every $\mu_2 > \mu_1$, i.e., $p_2 = \Phi(-\mu_2) < p_1 = \Phi(-\mu_1)$,

$$\frac{f(\underset{\sim}{Y}\mid\mu_2)}{f(\underset{\sim}{Y}\mid\mu_1)} = \left\{\frac{p_2}{p_1}\left(\frac{1-p_1}{1-p_2}\right)\right\}^T \left(\frac{1-p_2}{1-p_1}\right)^n,$$

which is nonincreasing in T.
By the Karlin-Rubin Theorem, a size-α UMP test is given by

$$\delta(\underset{\sim}{Y}) = \begin{cases} 1 & T(\underset{\sim}{Y}) < c \\ \gamma & T(\underset{\sim}{Y}) = c \\ 0 & T(\underset{\sim}{Y}) > c \end{cases}.$$

where $\gamma \in [0,1]$ and $c \in \mathbb{Z}^+$ satisfy $E_{\mu_0}\{\delta(\underline{Y})\} = \alpha$

That is, $P_{\mu_0}(T < c) + \gamma P_{\mu_0}(T = c) = \alpha$, assuming $T \sim \text{Binomial}(n, p_0)$ and $p_0 = \Phi(-\mu_0)$.

More specifically, one first finds a c such that

$$P_{\mu_0}(T < c) = \sum_{t=0}^{c-1} \binom{n}{t} p_0^t (1-p_0)^{n-t} \leq \alpha \quad \text{and}$$

$P_{\mu_0}(T < c+1) > \alpha$. Once c is chosen, then set

$$\gamma = \frac{\alpha - P_{\mu_0}(T < c)}{P_{\mu_0}(T = c)}.$$

(d) One can propose a C.I for p first, then convert it to an interval for μ using the relationship $\mu = -\Phi^{-1}(p)$

DAY #1, Q2 .

1. Infectious diseases are sometimes modelled with a so called *SIR* model (the letters stand for Susceptible, Infected, and Recovered). People begin in class *S*, then possibly migrate to class *I* (i.e., become infected), and then to class *R* (i.e., recover); no other transitions are possible. In a simple version of the model, the i^{th} individual begins in class S, waits a random amount of time $T_i \sim Exp(1/\lambda)$ before migrating to class *I*, then waits another random amount of time $U_i \sim Exp(1/\mu)$ before migrating to class R, with all the exponentially-distributed random variables T_i and U_i independent. Here a random variable $X \sim Exp(1/\lambda)$ if X has the p.d.f. $p(x) = \lambda \exp(-\lambda x)$ for $x > 0$.

   (a) For each $t > 0$, find the C.D.F. $Pr(T_i \leq t)$.

   $Pr(T_i \leq t) = \int_0^t \lambda \exp\{-\lambda x\} dx = 1 - \exp\{-\lambda t\}$

   (b) Let N denote the number of Susceptibles at time 0 and let X_t be the number of these who become infected by time t. Find the probability distribution of X_t.

   Because each T_i is independent with constant "success" probability given by (a), we know $X_t \sim Bin(N, 1 - \exp\{-\lambda t\})$.

   (c) Let W_1 be the length of time until the *first* of those people becomes infected. Find the probability distribution for W_1.

   Recall the c.d.f. of the minimum of independent random variables is given by,

   $$F_{W_1}(t) = 1 - [1 - F_{T_i}(t)]^N$$

   where $F_{T_i}(t)$ is the c.d.f of T_i. Therefore, the probability distribution for W_1 is given by

   $$F_{W_1}(t) = 1 - [1 - (1 - \exp\{-\lambda t)\}]^N = 1 - \exp\{-\lambda N t\} \sim Exp(N\lambda)$$

   (d) Let W_N be the length of time until the *last* of those people becomes infected. Find the probability density function for W_N.

   Recall the c.d.f. of a maximum of independent random variables is given by,

   $$F_{W_N}(t) = [F_{T_i}(t)]^N = [1 - \exp\{-\lambda t\}]^N,$$

   Now, differentiating yields the density function,

   $$f_{W_N}(t) = \lambda N [1 - \exp\{-\lambda t\}]^{N-1} \exp\{-\lambda t\}.$$

   (e) Let $Y_i = T_i + U_i$ be the total amount of time the i^{th} Susceptible waits before joining class R. Find the probability distribution of Y_i under the (simplifying) assumption $\lambda = \mu$.

Consider the moment generating function of Y_i,

$$\mathrm{E}(e^{tY_i}) = \mathrm{E}(e^{t(T_i+U_i)}) = \mathrm{E}(e^{tT_i})\mathrm{E}(e^{tU_i}) = (1 - t/\lambda)^{-2}$$

which is the moment generating function of a $Ga(2, \lambda)$ random variable.

2. *More Bang!*

Investors are concerned about the return of their money. Suppose you invested \$1000 in the US stock market last year. After one year, your investment is worth \$1080. The return is then $1080/1000 = 1.08$, and the rate of the return is $\log(1080/1000) = 0.077$.

Now, we consider a model for investment strategy. Label the initial value of of your investment as W_0,(e.g. $W_0 = \$1000$) and the annual return of the investment as R_t (e.g. $R_1 = 1.08$) during year t. The value at the end of the first year is $W_1 = W_0 R_1$, and by the end of year T the value is

$$W_T = W_{T-1}R_T = \cdots = W_0 R_1 R_2 \cdots R_T.$$

We suppose $\{R_t\}_{t=0}^T$ are i.i.d. random variables.

Based upon the historical data, we have summary statistics of the return R_t for different assets as in Table 1.

	Stocks	T-bills
Mean	1.10	1.05
Std Dev	0.20	0.04

Table 1: *Mean, standard deviations of annual returns, R_t, on US stocks and Treasury Bills*

If we start with \$1000 in each of the stock and the T-bills, we would expect to have \$1,100 in stock and \$1,050 in T-bills after one year. Because the expected value of a product of independent random variables is the product of expectations, we can find the expectations for each investment over a longer horizon given this assumption. Over 20 years, the initial investment of \$1000 in stock grows in expectation to $\$1000 \times (1.1)^{20} = \$6727$. By comparison, the initial investment in T-bills grows to \$2653.

At first glance, the above calculation of expected values seems quite reasonable. However, it uses only the mean of the returns and has no appreciation of the standard deviation - the risks!

To have a deeper understanding of impact of the variance on the long term returns, we convert the product to a sum:

$$\log(W_T) = \log(W_0) + \sum_{t=1}^{T} \log(R_t) = \log(W_0) + \sum_{t=1}^{T} r_t$$

where $r_t = \log(R_t)$ is the *continuously compounded rate of return* for $t = 1, \cdots, T$. Now for large T and by law of large numbers, we have

$$\log(W_T) \approx \log(W_0) + T\mathrm{E}(r_t)$$

(a) Let μ_r be the expectation of the rate of return (i.e. $\mu_r \equiv \mathrm{E}(r_t)$). Approximate μ_r for stocks and T-bills, respectively, using the mean and variance for R_t in Table 1 and the second-order Talor

DAY 1
Q3

IOS Problem

(a) $IOS = \log \prod_{i=1}^{n} f(Y_i|\hat{\theta}) - \log \prod_{i=1}^{n} f(Y_i|\hat{\theta}_{(i)})$

$= \sum_{i=1}^{n} \log f(Y_i|\hat{\theta}) - \sum_{i=1}^{n} \log f(Y_i|\hat{\theta}_{(i)})$

$= \sum_{i=1}^{n} \{\log f(Y_i|\hat{\theta}) - \log f(Y_i|\hat{\theta}_{(i)})\}$

$= \sum_{i=1}^{n} \{ l(Y_i; \hat{\theta}) - l(Y_i; \hat{\theta}_{(i)}) \}.$

(b) It is straightforward to show that

$\hat{\theta} = \bar{Y}$ (under Poisson assump.)

$\rightarrow \hat{\theta}_{(i)} = \bar{Y}_{(i)},$

where

$\bar{Y}_{(i)} = \frac{1}{n-1}(Y_1 + \dots + Y_{i-1} + Y_{i+1} + \dots + Y_n)$

Here,

$f(y|\theta) = \frac{\theta^y e^{-\theta}}{y!}$

$\rightarrow \log f(y|\theta) = y \log \theta - \theta - \log y!$

Therefore

$$\ell(Y_i;\hat{\theta}) = Y_i \log \hat{\theta} - \hat{\theta} - \log Y_i!$$
$$\ell(Y_i;\hat{\theta}_{(i)}) = Y_i \log \hat{\theta}_{(i)} - \hat{\theta}_{(i)} - \log Y_i!$$

and

$$\begin{aligned} IOS &= \sum_{i=1}^{n} \Big\{ Y_i \log \hat{\theta} - \hat{\theta} - \cancel{\log Y_i!} \\ &\qquad - Y_i \log \hat{\theta}_{(i)} + \hat{\theta}_{(i)} + \cancel{\log Y_i!} \Big\} \\ &= \sum_{i=1}^{n} (\hat{\theta}_{(i)} - \hat{\theta}) + \sum_{i=1}^{n} Y_i (\log \hat{\theta} - \log \hat{\theta}_{(i)}) \\ &= \sum_{i=1}^{n} (\overline{Y}_{(i)} - \overline{Y}) + \sum_{i=1}^{n} Y_i \log \left(\frac{\overline{Y}}{\overline{Y}_{(i)}} \right). \end{aligned}$$

Now, to show the approximate equality. First prove

$$\overline{Y}_{(i)} - \overline{Y} = \frac{1}{n-1} (\overline{Y} - Y_i).$$

Proof:

$$\begin{aligned} \overline{Y}_{(i)} - \overline{Y} &= \frac{1}{n-1} (Y_1 + \dots + Y_{i-1} + Y_{i+1} + \dots + Y_n) - \overline{Y} \\ &= \frac{1}{n-1} \left[\left(\sum_{i=1}^{n} Y_i - Y_i \right) - (n-1) \overline{Y} \right] \end{aligned}$$

$$= \frac{1}{n-1}\left[\cancel{n\bar{Y}} - Y_i - \cancel{n\bar{Y}} + \bar{Y} \right]$$

$$= \frac{1}{n-1}\left(\bar{Y} - Y_i \right),$$

as claimed.

Now,

$$IOS = \sum_{i=1}^{n} \left(\bar{Y}_{(i)} - \bar{Y} \right) + \sum_{i=1}^{n} Y_i \log\left(\bar{Y} / \bar{Y}_{(i)} \right)$$

$$= \underbrace{\sum_{i=1}^{n} (n-1)^{-1} \left(\bar{Y} - Y_i \right)}_{=0} + \sum_{i=1}^{n} Y_i \left\{ \log \bar{Y} - \log \bar{Y}_{(i)} \right\}$$

$$\approx \sum_{i=1}^{n} Y_i \left\{ \log \bar{Y} - \underbrace{\left((\log \bar{Y}) - \left(\bar{Y}_{(i)} - \bar{Y} \right) / \bar{Y} \right.}_{\text{TS expansion.}} \right\}$$

$$= -\sum_{i=1}^{n} Y_i \left(\bar{Y}_{(i)} - \bar{Y} \right) / \bar{Y}$$

$$= -\sum_{i=1}^{n} Y_i (n-1)^{-1} \left(\bar{Y} - Y_i \right) / \bar{Y}$$

$$= \frac{1}{(n-1)\bar{Y}} \sum_{i=1}^{n} Y_i \left(Y_i - \bar{Y} \right)$$

$$= \frac{1}{(n-1)\bar{Y}} \sum_{i=1}^{n} \left(Y_i - \bar{Y} \right)\left(Y_i - \bar{Y} \right)$$

4

$$= \frac{1}{(n-1)\bar{Y}} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \frac{S^2}{\bar{Y}}.$$

(c) By the WLLN,

$$\bar{Y} \xrightarrow{p} \theta \quad \text{(population mean)}$$
$$S^2 \xrightarrow{p} \theta \quad \text{(population variance)}$$

By continuity,

$$\frac{S^2}{\bar{Y}} \xrightarrow{p} \frac{\theta}{\theta} = 1, \quad \text{as } n \to \infty$$

Note: $g(s,t) = \frac{s}{t}$ is a continuous function for $s > 0$, $t > 0$.

(d) A sensible test statistic would be

$$T = \left| \frac{S^2}{\bar{Y}} - 1 \right|.$$

Deriving the distribution of T under

H_0: Poisson (θ) model holds

would be quite difficult. Could use a bootstrap or other simulation procedure.