Sample Questions for Module 1

- Module 1 is about estimation.
- Main concepts: bias, MSE, CI, and hypothesis tests
- Main skills: calculate bias and MSE of an estimator, derive CI and hypothesis tests, calculate and interpret Type I and Type II errors, power, and sample size

Sample Question 1

Suppose $Y_1, Y_2, \dots, Y_n$ are iid Poisson random variables with rate $\lambda > 0$, and that $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is to be used as an estimator for λ.

(a) Show that $\bar{Y}$ is an unbiased estimator for λ.

(b) Calculate the MSE of $\bar{Y}$.

$E(Y_i) = \lambda$, $var(Y_i) = \lambda$

(a) $$E(\bar{Y}) = E\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right] = \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = \frac{1}{n}\sum_{i=1}^{n}(\lambda) = \frac{n\lambda}{n} = \lambda$$

As $E(\bar{Y}) = \lambda$, $\bar{Y}$ is an unbiased estimator for λ.

Sample Question 1 Solution

b) $MSE_{\bar{Y}}(\lambda) = var(\bar{Y}) + b_{\bar{Y}}(\lambda)^2$

$= var(\bar{Y})$ (as $\bar{Y}$ is unbiased, $b_{\bar{Y}}(\lambda) = 0$)

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

$= \frac{1}{n^2}\sum_{i=1}^{n} var(Y_i)$ (by independence of Y_i)

$= \frac{1}{n^2}(n\lambda)$

$= \frac{\lambda}{n}$

Sample Question 2

An electronics manufacturer observed that, on average, 3% of its top-selling product failed quality control. The engineers redesigned the product aiming to improve this fail rate. An initial test batch of 100 samples were produced. If no or only one of these samples failed quality control, then the company will approve the redesigned product for production. Let p denote the true proportion of the redesigned product failing quality control.

(a) Identify the null and alternative hypotheses if we were to test whether the redesigned product will have an improved failure rate.

(b) Describe what the Type I and Type II errors are in the context of this problem.

(c) If the true value of p is 0.01, what is the probability that the company will commit a Type II error?

Sample Question 2 Solution

a) $H_0: p = 0.03$ vs $H_a: p < 0.03$

b) Type I error:
- reject H_0 when H_0 is true
- none or only one of the 100 samples failed quality control, when the true proportion of failure is 0.03
- company went ahead with production of the redesigned product when in fact its proportion of failure is the same as the current version of the product

Type II error:
- fail to reject H_0 when it is false
- two or more of the 100 samples failured quality control, but the true proportion of failure was lower than 0.03
- company did not go ahead with production of the redesigned product when it should have

Sample Question 2 Solution

c) Let Y = number of samples in the 100 samples fail quality control

$$Y \sim B(100, p)$$

$$
\begin{aligned}
P(\text{Type II error}) &= P(\text{not reject } H_0 \mid H_0 \text{ false}) \\
&= P(Y \geq 2 \mid p = 0.01) \\
&= 1 - P(Y < 2 \mid p = 0.01) \\
&= 1 - P(Y = 0 \mid p = 0.01) - P(Y = 1 \mid p = 0.01) \\
&= 1 - \underbrace{\binom{100}{0} 0.01^{0}\, 0.99^{100}}_{\text{dbinom}(0,100,0.01)} - \underbrace{\binom{100}{1} 0.01^{1}\, 0.99^{99}}_{\text{dbinom}(1,100,0.01)} \\
&= 1 - 0.3660 - 0.3697 \\
&= 0.264
\end{aligned}
$$

Sample Question 3

Suppose $Y_1, Y_2, \ldots, Y_n$ are iid $N(\mu, \sigma^2)$ random variables with σ^2 known.

(a) Carefully define a $100(1-\alpha)\%$ confidence interval.

(b) State the distribution of $\bar{Y}$.

(c) Find the distribution of $Z = \frac{\bar{Y}-\mu}{\sigma/\sqrt{n}}$ using the method of moment generating function. You may assume that if $Y \sim N(\mu, \sigma^2)$, then $M_Y(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}$.

(d) Using Z as a pivotal quantity, construct a $100(1-\alpha)\%$ confidence interval for μ.

Sample Question 3 Solution

(a) A $100(1-\alpha)\%$ confidence interval for θ is a random interval (L, U) that satisfies $P(L < \theta < U) = 1-\alpha$.

(b) $\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

(c) $M_{\bar{Y}}(t) = e^{\mu t + \frac{1}{2}t^2\left(\frac{\sigma^2}{n}\right)}$

$$M_Z(t) = E[e^{tZ}]$$
$$= E\left[e^{t\left(\frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)}\right]$$
$$= e^{-\frac{t\mu}{\frac{\sigma}{\sqrt{n}}}} E\left[e^{\boxed{\frac{t\bar{Y}}{\frac{\sigma}{\sqrt{n}}}}}\right]$$
$$= e^{-\frac{t\mu}{\frac{\sigma}{\sqrt{n}}}} M_{\bar{Y}}\left(\frac{t}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= e^{\cancel{\frac{-t\mu}{\frac{\sigma}{\sqrt{n}}}}} e^{\cancel{\mu\left(\frac{t}{\frac{\sigma}{\sqrt{n}}}\right)} + \frac{1}{2}\left(\frac{t}{\cancel{\frac{\sigma}{\sqrt{n}}}}\right)^2\left(\cancel{\frac{\sigma^2}{n}}\right)}$$
$$= e^{\frac{1}{2}t^2}$$

which is the MGF of a $N(0,1)$ random variable.

So $Z \sim N(0,1)$.

Sample Question 3 Solution

(d) $1-\alpha = P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}})$

$$= P\left(-Z_{\frac{\alpha}{2}} < \frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\frac{\alpha}{2}}\right)$$

$$= P\left(-Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \bar{Y}-\mu < Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(-Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu-\bar{Y} < Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(\bar{Y}-Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{Y}+Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

The $100(1-\alpha)\%$ confidence interval for μ is

$$\left(\bar{Y}-Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},\ \bar{Y}+Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

Or can also write it as $\bar{Y} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.