SOLUTIONS TO PRACTICALS 7 & 8

PRAC. 8 Q1

Let $N(t)$ denote the number of potholes occurring in $100t$ metres. Assuming that potholes occur randomly at a constant rate per 100m, then $N(t) \sim Po(\lambda t)$.

We test $H_0: \lambda = 6$ vs $H_1: \lambda > 6$.

Assume H_0 is true. Then $N(t) \sim Po(6t)$, and so the number of potholes in 500m is distributed as $N(5) \sim Po(30)$.

We calculate $p_0^+ = P(N(5) \geq 40) = P(Po(30) \geq 40) = 0.0463$

The test is significant at the 5% level; moderate evidence against H_0 in favour of H_1. We would therefore resurface the road, since we have moderate evidence that the pothole rate is greater than the acceptable 6 per 100m.

PRACTICAL 8 Q2

Assume that assaults occur randomly at a constant rate of λ assaults per week. Then, letting $N(t)$ denote the number of assaults in t weeks, $N(t) \sim Po(\lambda t)$. $t = 24$ in this case.

The estimated rate of assaults is $\hat{\lambda} = \frac{32}{24} = \frac{4}{3}$, while

$\mathrm{Var}(\hat{\lambda}) = \frac{\lambda t}{t^2} = \frac{\lambda}{t}$, so $ESE(\hat{\lambda}) = \sqrt{\frac{\hat{\lambda}}{t}} = \sqrt{\frac{(4/3)}{24}} = 0.2357$.

For a 95% confidence interval, $\alpha = \frac{1}{2}(1 - 0.95) = 0.025$. Using $N(0,1)$ tables, $z_{0.025} = 1.9600$. It follows that the 95% confidence interval is given by $\frac{4}{3} \pm z_{0.025} \times ESE(\hat{\lambda}) = 1.333 \pm 1.96 \times 0.2357$

$= 1.333 \pm 0.461972$,

So the 95% CI for λ is $(0.8714, 1.7953)$.

PRACTICAL 7 Q1

Assuming matches are independent with Roy having a fixed probability of scoring, p, the number of matches out of 10 in which Roy scores is distributed as $R \sim Bin(10, p)$.

We test $H_0: p = 0.75$ vs $H_1: p < 0.75$.

Under H_0, $R \sim Bin(10, 0.75)$.

We calculate $p_0^- = P(R \leq 2) = P(Bin(10, 0.75) \leq 2)$ (0, 1, 2 successes = 10, 9, 8 failures)

$= P(Bin(10, 0.25) \geq 8)$

$= 0.0004$ (from tables).

The test is significant at the 0.1% level. Very strong evidence against H_0 in favour of H_1. Roy has indeed lost his touch.

PRACTICAL 7 Q2

Assuming households' viewing is independent with a fixed probability p of watching the show, the number of households out of n that watch the show, $R \sim \text{Bin}(n, p)$. $n = 443$ in this case

We test $H_0: p = 0.25$ vs $H_1: p < 0.25$.

Under H_0,

$$p_0^- = P(R \leq 102) = P(\text{Bin}(443, 0.25) \leq 102)$$
$$\doteqdot P(N(110.75, 83.0625) < 102.5)$$
$$= P\left(N(0,1) < \frac{102.5 - 110.75}{\sqrt{83.0625}}\right)$$
$$= P(N(0,1) < -0.90521)$$
$$= P(N(0,1) > 0.90521) \text{ by symmetry}$$
$$\approx 0.1814 \quad \text{(from tables)}$$

The test is not significant at the 10% level; insufficient evidence to reject H_0. Celebrity statistics gets to live another day!

PRACTICAL 7 Q3

(i) Since $n = 185$ is large and $p = 0.4$ is not too close to 0 or 1,

$$\text{Bin}(185, 0.4) \doteqdot N(185 \times 0.4, 185 \times 0.4 \times 0.6) = N(74, 44.4)$$

(ii) Since $n = 360$ is large and $p = 0.05$ is small,

$$\text{Bin}(360, 0.05) \doteqdot Po(360 \times 0.05) = Po(18).$$

PRACTICAL 8 Q3

See corresponding lecture which contained a very similar example.

PRACTICAL 8 Q4

H_0 states that Y has the Poisson distribution, with p.m.f.

$$P(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \ldots.$$

Because λ is unknown, we approximate it as

$$\hat{\lambda} = \frac{N(75)}{75} = \frac{36}{75} = 0.48 \text{ accidents per week.}$$

Under H_0:

$$\hat{p}_0 = P(Y = 0) = \frac{\lambda^0 e^{-\lambda}}{0!}, \quad \hat{p}_1 = \frac{\lambda^1 e^{-\lambda}}{1!}, \quad \hat{p}_2 = 1 - \hat{p}_0 - \hat{p}_1.$$

These probabilities are estimated by replacing λ with $\hat{\lambda}$, giving

$$\hat{p}_0 = 0.619, \quad \hat{p}_1 = 0.297, \quad \hat{p}_2 = 0.084.$$

The expected cell frequencies are therefore:

$$\mathbb{E}[n_0] = 46.409, \quad \mathbb{E}[n_1] = 22.276, \quad \mathbb{E}[n_2] = 6.315.$$

The χ^2 statistic has an approximate χ^2 distribution with $3-1-1 = 1$ degree of freedom.

$$\chi^2_{obs} = 2.017$$

From tables, $P(\chi^2_{(1)} > \chi_{obs})$ lies between 0.1 and 0.2 (exact value from R is 0.1555). Therefore do not reject H_0. The data does not present sufficient evidence to contradict our hypothesis that Y possesses a Poisson distribution.