UNIVERSITY OF BRISTOL

School of Mathematics

ANOMALY DETECTION — solutions
MATHM0028
[Paper code MATH–M0030J (Mock)]

January 2019 1 hours 30 minutes (this mock: 45 minutes)

1. (a) (**Bookwork**) A composite hypothesis is a hypothesis in which not all the parameters are specified. (**2 marks**)

   (b) (**Bookwork**) The Bonferroni correction procedure rejects each hypothesis i for which $p_i \leq \alpha/n$. (**3 marks**) Assume the true hypotheses are $H_0^{(1)}, \ldots, H_0^{(n_0)}$, for $n_0 \leq n$. By Boole's inequality,

$$\mathbb{P}\left[\bigcup_{i=1}^{n_0}\left\{P_i \leq \frac{\alpha}{n}\right\}\right] \leq \sum_{i=1}^{n_0} \mathbb{P}(P_i \leq \alpha/n) = \frac{n_0}{n}\alpha \leq \alpha.$$

   (**7 marks**)

   (c) i. (**Seen similar**) We have:

$$\begin{aligned}
\mathbb{P}(Y_{(k)} \leq y) &= \mathbb{P}(\text{at least } k \text{ of } Y_1, \ldots, Y_n \leq y) \\
&= \mathbb{P}\left\{\sum_{i=1}^{n} \mathbb{I}(Y_i \leq y) \geq k\right\} \\
&= \mathbb{P}\left\{\sum_{i=1}^{n} \mathbb{I}(Y_i \leq y) > k-1\right\} \\
&= 1 - \mathbb{P}\left\{\sum_{i=1}^{n} \mathbb{I}(Y_i \leq y) \leq k-1\right\} \\
&= 1 - F_{\text{Binomial}}\{k-1, F_{\text{Beta}}(y, \alpha, \beta), n\}.
\end{aligned}$$

   (**7 marks**)

   ii. (**Unseen and moderately challenging**) We know that the minimum of m independent uniform random variables on $[0, 1]$ is distributed as a $\text{Beta}(1, m)$ variable. Therefore, $W_1, \ldots, W_n$ are independent $\text{Beta}(1, m)$ variables. By the previous question, we therefore have

$$\mathbb{P}(W_{(5)} \leq w) = 1 - F_{\text{Binomial}}\{4, F_{\text{Beta}}(w, 1, m), n\}.$$

   (**7 marks**)

iii. (**Unseen and moderately challenging**) Since we are interested in small values of $w_{(5)}$, the combined p-value is

$$1 - F_{\text{Binomial}}\{4, F_{\text{Beta}}(10^{-5}, 1, 100), 50\}.$$

Since F_{Binomial} is a cumulative distribution function it is non-decreasing, and therefore

$$F_{\text{Binomial}}\{4, F_{\text{Beta}}(10^{-5}, 1, 100), 50\} \geq F_{\text{Binomial}}\{1, F_{\text{Beta}}(10^{-5}, 1, 100), 50\}.$$

Therefore,

$$\begin{aligned} 1 - F_{\text{Binomial}}\{4, F_{\text{Beta}}(10^{-5}, 1, 100), 50\} &\leq 1 - F_{\text{Binomial}}\{1, F_{\text{Beta}}(10^{-5}, 1, 100), 50\} \\ &\leq 0.01, \end{aligned}$$

so that we can reject the null hypothesis of no infection at a false positive rate of $0.05 \geq 0.01$.(**7 marks**)

(d) i. (**Bookwork**) We have:

$$\begin{aligned} p\{\mu \mid N(\cdot)\} &= p\{\mu \mid A(\cdot)\} \\ &\propto p\{A(\cdot) \mid \mu\} p(\mu) \\ &\propto \mu^{A(T)}(1-\mu)^{T-A(T)} \mu^{\alpha_\mu - 1}(1-\mu)^{\beta_\mu - 1} \\ &\propto \mu^{\{A(T)+\alpha_\mu\}-1}(1-\mu)^{\{T-A(T)+\beta_\mu\}-1}, \end{aligned}$$

from which we recognise the functional form of a Beta$\{\alpha_\mu + A(T), \beta_\mu + T - A(T)\}$ distribution. (**5 marks**)

ii. (**Seen similar, but moderately challenging**) We have:

$$\begin{aligned} p\{\nu \mid N(\cdot)\} &= p\{\nu \mid (\mathrm{d}B(t) : \mathrm{d}A(t) = 1)\} \\ &\propto p\{(\mathrm{d}B(t) : \mathrm{d}A(t) = 1) \mid \nu\} p(\nu) \\ &\propto \nu^{N(T)-A(T)}(1-\nu)^{2A(T)-N(T)} \nu^{\alpha_\nu - 1}(1-\nu)^{\beta_\nu - 1} \\ &\propto \nu^{\{N(T)-A(T)+\alpha_\nu\}-1}(1-\mu)^{\{2A(T)-N(T)+\beta_\nu\}-1}, \end{aligned}$$

from which we recognise the functional form of a Beta$[\alpha_\nu + N(T) - A(T), \beta_\nu + 2A(T) - N(T)]$ distribution.(**6 marks**)

iii. (**Seen similar**) We have

$$\mathbb{P}\{N(T+1) = 2 \mid N(\cdot)\} = \mathbb{P}\{\mathrm{d}A(T+1) = 1 \cap \mathrm{d}B(T+1) = 1 \mid N(\cdot)\}.$$

Conditional on $N(\cdot)$, $\mathrm{d}A(T+1)$ and $\mathrm{d}B(T+1)$ are independent Bernoulli variables with success probabilities $\{A(T) + \alpha_\mu\}/\{T + \alpha_\mu + \beta_\mu\}$ and $\{N(T) - A(T) + \alpha_\nu\}/\{A(T) + \alpha_\nu + \beta_\nu\}$, so that:

$$\mathbb{P}\{N(T+1) = 2 \mid N(\cdot)\} = \frac{\{A(T) + \alpha_\mu\}\{N(T) - A(T) + \alpha_\nu\}}{\{T + \alpha_\mu + \beta_\mu\}\{A(T) + \alpha_\nu + \beta_\nu\}}.$$

(**6 marks**)