RSS Higher Certificate in Statistics, Specimen A

Module 3: Basic Statistical Methods

Solutions

Question 1

(i)

帰無仮説: 200C° と 250C° において鉄鋼の破壊応力の母平均には違いはない.

対立仮説: 破壊応力の母平均には違いがあり， 250C° の方ときの方が大きい.

$n_1 = 8, n_2 = 7, \bar{x}_1 = 59.6, \bar{x}_2 = 63.6, s_1^2 = 17.4^2, s_2^2 = 20.1^2$ である．分散が共通であると仮定する（ (ii) を見よ）とプールされた分散の推定値は

$$s^2 = \left(\left(7 \times 17.4^2\right) + \left(6 \times 20.1^2\right)\right) / 13 = 349.91 \text{（すなわち } s = 18.69\text{）}$$

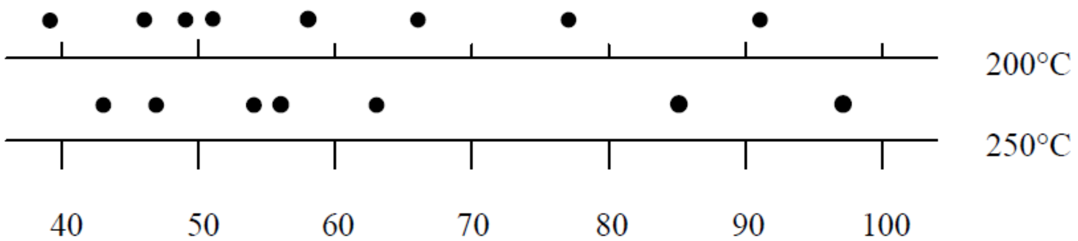
であり，自由度は 13 である．したがって，母平均が同じであるという仮説を検定するための検定統計量の値は

$$\frac{\bar{x}_2 - \bar{x}_1 (-0)}{s\sqrt{\frac{1}{8} + \frac{1}{7}}} = \frac{4.0}{9.681} = 0.41$$

であり，これを t_{13} と比較する．これは有意水準 5%で有意ではない（片側 5%点は 1.771）ので，帰無仮説を棄却する根拠はない．すなわち母平均は同じである．

(ii)

2 つの母集団は同じ分散をもつ正規分布であると仮定する．下のドットプロットはこれらの仮定をチェックするのに役立つ．



サンプルサイズは小さいが，範囲は似ており，似たような分散を仮定するのに合理的であろう．しかしながら，分布の中心に観測値が集まっていなかったり，いくつかゆがみやはずれ値が認められるので，正規性の仮定についてはいくつか疑問がある．

(iii)

ウィルコクソンの順位和検定が適している.（マン・ホイットニー検定でも本質的に同じである.）帰無仮説は250C°の母集団の中央値が200C°のものと等しいというものであり，対立仮説は前者の方が大きいというものである．まず以下のように全ての15個のデータをランク付けする．

Data	39	43	46	47	49	51	54	56	58	63	66	77	85	91	97
Ranks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	A	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*A*	*B*

A refers to 200°C, *B* to 250°C.

小さい方のサンプルの順位和，つまり *B* の順位和は 2+4+7+8+10+13+15=59 となる．有意水準には再び 5% を用いる．ここでの検定は片側検定となる．対立仮説は「*B* の母集団の中央値が *A* のものよりも大きい」なので，*B* の順位和を検定統計量の帰無分布の上側 5%点と比較したい．ところが，試験で用いる王立統計協会の数表には，他で出版されている数表と同様に下側の点しか与えられていない．

これを克服する方法の一つは降順にデータをランク付けし直すことである．すなわち，高いものから低いものに並び直すということである．それにより下側 5%点と比較することが可能となる．また，数表より上側点を簡単に求めることができ，その方がより一般的な方法と言える．サンプルサイズが n_1 と n_2 で，統計量を求めたときのサンプルサイズが n_1 であるとき，帰無分布における統計量の平均は $n_1(n_1+n_2+1)/2$ である．帰無分布が平均について対称であることは明らかなので，もし下側点が平均を *a* だけ下回る点であるならば，対応する上側点は平均を *a* だけ上回る点となる．ここで $(n_1, n_2) = (7, 8)$ の場合の下側 5%点は 41 である．平均は $7\times16/2=56$ であるので上側 5%点は 71 でとなる．したがってサンプルから計算された統計量の値 59 は上側 5%の棄却域の外にあるので，帰無仮説を棄却する根拠はない．2 つの母集団はこの点では同じであるように見える．

(iv)

どちらの検定も破壊応力が増加するという仮説を支持はしていない．また，*t* 検定の根拠となる正規性の仮定を考慮するには疑問があるので，ノンパラメトリック検定の方が適していると言える．

Question 2

(i)

2×2 の分割表を用いる．帰無仮説は性別とその人が最近コレステロール値を測定したかどうかとの間に関係がないというものである．分割表は以下のようになっており，各セルのカッコの中が期待頻度となっている（例：$88.48 = 131 \times 206 / 305$）.

		Cholesterol level recorded		
		No	*Yes*	Total
Sex	*Female*	109 (88.48)	22 (42.52)	131
	Male	97 (117.52)	77 (56.48)	174
	Total	206	99	305

観測された頻度と期待頻度との間の全ての差は ± 20.52 であり，Yates の補正を用いるのならば ± 20.02 となる．したがって検定統計量は（Yates の補正を用いて）次のように計算することができる．

$$(20.02)^2 \left\{ \frac{1}{88.48} + \frac{1}{42.52} + \frac{1}{117.52} + \frac{1}{56.48} \right\} = 24.46$$

（Yates の補正を用いなければ 25.70 となる．）この値を χ_1^2 と比較する．これは非常に高有意である（例えば，1%点は 6.635 である）．帰無仮説を棄却するのに非常に強い根拠があり，関連があると結論付けられる．

(ii)

$p_f - p_m$ は $\hat{p}_f - \hat{p}_m = \frac{22}{131} - \frac{77}{174} = 0.168 - 0.443 = -0.275$ と推定される．また $\hat{p}_f - \hat{p}_m$ の分散の推定値は $\frac{\hat{p}_f(1-\hat{p}_f)}{n_f} + \frac{\hat{p}_m(1-\hat{p}_m)}{n_m} = 0.001067 + 0.001418 = 0.002485$ で与えられる．よって，$p_f - p_m$ の近似 95%信頼区間は $-0.275 \pm \left(1.96 \times \sqrt{0.002485}\right) = (-0.177, -0.373)$ となる．

(iii)

男性と女性で割合が同じでないということに明らかな根拠がある．(i) では，これは関連の強い根拠があることによって説明される．(ii) では信頼区間が 0 を含まず，さらに 0 からとても離れているので，(i) と同様に割合が実際に違うというとても強い根拠を与えてくれる．

Question 3

(i)

(a) 第一種の過誤は実際には帰無仮説が正しいにもかかわらず，帰無仮説を棄却し，対立仮説を採択してしまうことである．

(b) 第二種の過誤は実際には対立仮説が正しいにもかかわらず，帰無仮説を棄却し損ねてしまうことである．

(c) 検定の有意水準は帰無仮説が正しい時，帰無仮説を棄却する確率である．すなわち第一種の過誤をおかしてしまう確率である．慣例としてα と表す.

(d) 検出力は帰無仮説が棄却される確率であり，パラメータの関数（あるいは一つのパラメータに対する検定でなくても）として表される．β を第二種の過誤をおかしてしまう確率とすると検出力は$1-\beta$ で与えられる．

(ii)

X を瓶の中にあるコーヒーの量とすると X ~ N(μ, 15^2) である．サンプルサイズは $n=9$ なので，$\overline{X}$ ~ N(μ, 15^2 /9) となる．以降，Z ~ N(0, 1) とする．

(a) $\mu = 200$ であるので，

$$P\left(\bar{X} < 190\right) = P\left(Z < \frac{190-200}{15/3} = -2.0\right) = 0.02275,$$

$$P\left(\bar{X} > 210\right) = P\left(Z > \frac{210-200}{15/3} = 2.0\right) = 0.02275.$$

したがって第一種の過誤をおかす確率は $0.02275 + 0.02275 = 0.0455$ となる．

(b) ここでは $\mu = 216$ であるので，

$P\left(\bar{X} < 190\right) = P\left(Z < \frac{190-216}{15/3} = -5.2\right)$ で，これは小数点以下数桁まで 0 となる．また，

$$P\left(\bar{X} > 210\right) = P\left(Z > \frac{210-216}{15/3} = -1.2\right) = 0.1151.$$

したがって，生産した瓶が受け入れられてしまう合計の確率は 0.1151 となる．（これはこの検定方式に対しての第二種の誤りをしてしまう確率であり，すなわち $\mu = 216$ における β の値である．したがって，実際に $\mu = 216$ であるときのこの検定方式の検出力は $1-0.1151 = 0.8849$ となる．）

Question4

(i) 前提となる仮説（すなわち帰無仮説）のもとでの観測された頻度(o)と期待頻度(e)が次の表に与えられている．

	AB	Ab	aB	ab	
o	773	231	238	59	(Total 1301)
e	731.81	243.94	243.94	81.31	(Total 1301)

検定統計量は

$$X^2 = \sum \frac{(o-e)^2}{e} = \frac{(773-731.81)^2}{731.81} + \ldots + \frac{(59-81.31)^2}{81.31}$$
$$= 2.318 + 0.686 + 0.145 + 6.121 = 9.27$$

であり，χ_3^2 と比較する．（ここでは推定されているパラメーターがないことに注意する.） これは 5%の水準で有意である（5%点は 7.815）．帰無仮説は有意水準 5%点で棄却される．つまり比率 9:3:3:1 をここで適用する根拠はない．検定統計量に主に寄与している部分は ab のセルからであることに注意したい．すなわち，ab の観測頻度が期待されたものよりも小さく，これが想定した比率と食い違う主要な原因となっている..

(ii) まず表のはじめ 3 つのセルのみについて考える．9:3:3 という帰無仮説のもとでの観測頻度と期待頻度は次のようになる．

	AB	Ab	aB	
o	773	231	238	(Total 1242)
e	745.2	248.4	248.4	(Total 1242)

検定統計量は

$$X^2 = \sum \frac{(o-e)^2}{e} = \frac{(773-745.2)^2}{745.2} + \frac{(231-248.4)^2}{248.4} + \frac{(238-248.4)^2}{248.4}$$
$$= 1.037 + 1.219 + 0.435 = 2.69$$

これを χ_2^2 と比較する．これは有意ではない（5%点は 5.991）．AB， Ab， aB に対して 9:3:3（すなわち 3:1:1）という帰無仮説を棄却する根拠はない．

それでは次のように最初の 3 つのセルを結合させ， ab のセルを再び導入しよう．

	Combined	ab	
o	1242	59	(Total 1301)
e	1219.69	81.31	(Total 1301)

統計量は

$$X^2 = \sum \frac{(o-e)^2}{e} = \frac{(1242-1219.69)}{1219.69} + \frac{(59-81.31)^2}{81.31} = 0.408 + 6.121 = 6.53$$

であり χ_1^2 と比較する．これは 5%の水準で有意である（5%点は 3.841）．帰無仮説は 5%点の水準で棄却される．したがって比率 15:1 を適用しないという根拠が認められる．
概して，主に *ab* のタイプが十分な頻度で起きていないということを結果が示している．なぜならば, このタイプは (i) と (ii) 両方で相違があり, どちらも *e* よりも *o* の方が小さいからである．したがって生存割合が低いという推測をより詳しく調べる必要がある．

(iii) *ab* のタイプの割合 p は $\hat{p}=\frac{59}{1301}=0.0453$ と推定される． $\hat{p}$ の分散の推定値は

$\frac{\hat{p}(1-\hat{p})}{n}=0.00003324$ で与えられる．

よって， p の近似 95%信頼区間は $0.0453 \pm \left(1.96 \times \sqrt{0.00003324}\right)$，すなわち $(0.0340, 0.0566)$ で与えられる．$1/16\,(=0.0625)$ がこの区間に入っていないことに注意したい．区間は完全にこの値よりも下になっている．したがって *ab* のタイプの真の割合は 1/16 よりも小さいという根拠が認められる．