例題3

母分布：未知、母分散σ：未知、母平均μ：未知、標本サイズ：大標本
正規分布表かエクセルを使ってください

ある県内の農園から収穫したルレクチェ120個を無作為抽出したところ、平均468g, 標準偏差26gであった。母平均の95%信頼区間を求めよ。

標本平均 $\bar{x} = 468$, 標本の標準偏差$s = 26$, n=120

母平均μの95%信頼区間を求める。

標準化変換 $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$　信頼度95% => z=1.96

$$\rightarrow \mu = \bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$$

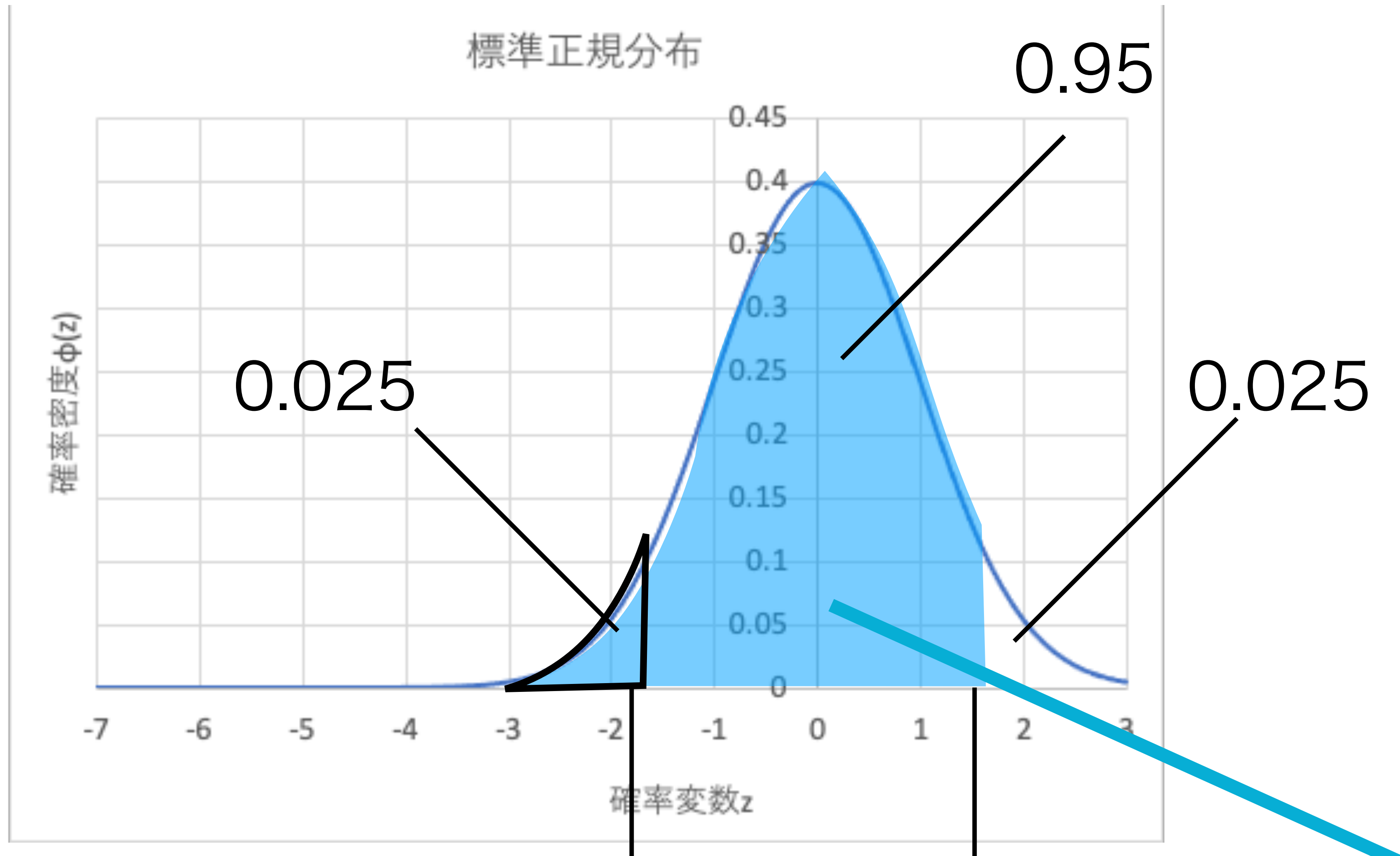
$$= \bar{x} \pm z \times \frac{1}{\sqrt{n}}\sqrt{\frac{n}{n-1}}s$$

$$= 468 \pm 1.96 \times \frac{1}{\sqrt{119}} \times 26$$

$463 \leq \mu \leq 473$ (g)

最後まで計算しましょう(小数点はok)

補足



累積確率が0.025+0.95=0.975のときのzを知りたい
正規分布表から、z=1.96

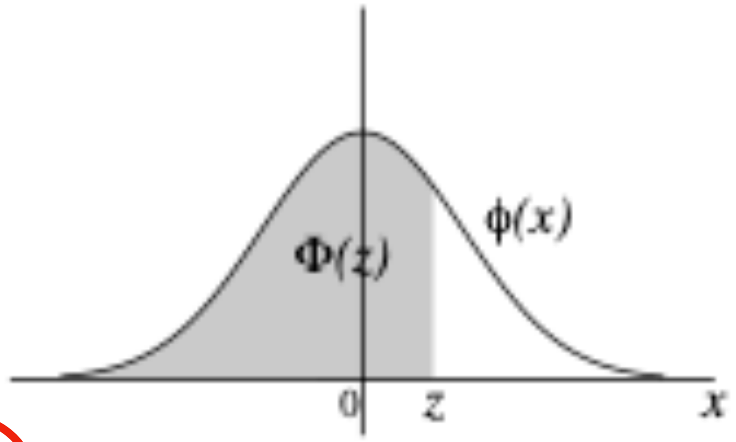
z=-1.96

(NORM.S.DIST(累積確率)
でもzが求まる)

B.2　正規分布表

正規分布表には積分範囲が異なるものがあるので注意すること．

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$



	0	1	2	3	4	5	6	7	8	9
0.0	.50000	.50398	.50797	.51196	.51595	.51993	.52392	.52790	.53188	.53585
0.1	.53982	.54379	.54775	.55171	.55567	.55961	.56355	.56749	.57142	.57534
0.2	.57925	.58316	.58706	.59095	.59483	.59870	.60256	.60641	.61026	.61409
0.3	.61791	.62171	.62551	.62930	.63307	.63683	.64057	.64430	.64802	.65173
0.4	.65542	.65909	.66275	.66640	.67003	.67364	.67724	.68082	.68438	.68793
0.5	.69146	.69497	.69846	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72574	.72906	.73237	.73565	.73891	.74215	.74537	.74857	.75174	.75490
0.7	.75803	.76114	.76423	.76730	.77035	.77337	.77637	.77935	.78230	.78523
0.8	.78814	.79102	.79389	.79673	.79954	.80233	.80510	.80784	.81057	.81326
0.9	.81593	.81858	.82121	.82381	.82639	.82894	.83147	.83397	.83645	.83891
1.0	.84134	.84375	.84613	.84849	.85083	.85314	.85542	.85769	.85992	.86214
1.1	.86433	.86650	.86864	.87076	.87285	.87492	.87697	.87899	.88099	.88297
1.2	.88493	.88686	.88876	.89065	.89251	.89435	.89616	.89795	.89972	.90147
1.3	.90319	.90490	.90658	.90824	.90987	.91149	.91308	.91465	.91620	.91773
1.4	.91924	.92073	.92219	.92364	.92506	.92647	.92785	.92921	.93056	.93188
1.5	.93319	.93447	.93574	.93699	.93821	.93942	.94062	.94179	.94294	.94408
1.6	.94520	.94630	.94738	.94845	.94949	.95052	.95154	.95254	.95352	.95448
1.7	.95543	.95636	.95728	.95818	.95907	.95994	.96079	.96163	.96246	.96327
1.8	.96406	.96485	.96562	.96637	.96711	.96784	.96855	.96925	.96994	.97062
1.9	.97128	.97193	.97257	.97319	.97381	.97441	.97500	.97558	.97614	.97670
2.0	.97724	.97778	.97830	.97882	.97932	.97981	.98030	.98077	.98123	.98169
2.1	.98213	.98257	.98299	.98341	.98382	.98422	.98461	.98499	.98537	.98573
2.2	.98609	.98644	.98679	.98712	.98745	.98777	.98808	.98839	.98869	.98898
2.3	.98927	.98955	.98982	.99009	.99035	.99061	.99086	.99110	.99134	.99157
2.4	.99180	.99202	.99223	.99245	.99265	.99285	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99429	.99445	.99461	.99476	.99491	.99505	.99520
2.6	.99533	.99547	.99560	.99573	.99585	.99597	.99609	.99620	.99631	.99642
2.7	.99653	.99663	.99673	.99683	.99692	.99702	.99710	.99719	.99728	.99736
2.8	.99744	.99752	.99759	.99767	.99774	.99781	.99788	.99794	.99801	.99807
2.9	.99813	.99819	.99824	.99830	.99835	.99841	.99846	.99851	.99855	.99860

(以下次ページ)

エクセルや正規分布表を使用のこと

例題4　テキスト問題8-2

遺跡から貝殻が12個出土したので、その質量のデータを使って現存種の貝と同種のものであるかどうかを検定したい。データは次の通り(単位はg)。

11.78, 12.92, 7.55, 14.52, 12.05, 19.0, 11.29, 11.81, 15.38, 9.62, 14.19, 12.62

これと比較したい現存種の貝の質量の平均値は12.6 g 標準偏差は1.9 gである。この母集団は正規分布しているものとする。

平均質量の分布を考えて検定を行ったとき、この貝殻は現存種のものである、という仮説を危険度5%で棄却できるか。

テキストappendix E解答を参照してください

例題5 統計検定２級2019.11.問18-1

次の表は、 2017年の2人以上の労働者世帯について、 47都道府県庁所在地市別に1世帯当たり1ヶ月間の収入と支出をまとめたものである(単位：万円)。なお、以下の表における世帯主収入の合計は、定期収入と賞与の和である。資料：総務省「2017年家計調査年報」

	世帯主収入			
	定期収入	賞与	合計	消費支出
札幌市	34.8	7.9	42.7	30.7
青森市	28.1	5.3	33.4	26.9
盛岡市	35.4	6.6	42	30.7
仙台市	30.6	5.3	35.9	30.9
…				
大分市	36.6	8.0	44.6	32.2
宮崎市	29.9	5.6	35.5	30.3
鹿児島市	33.5	6.4	39.9	30.9
那覇市	27.6	4.4	32	26.4

消費支出が定期収入及び賞与で説明できるかどうかを検証するため、次の重回帰モデルを考える。

消費支出=$\alpha_0 + \alpha_1 \times$ 定期収入 $+\alpha_2 \times$ 賞与$+u$

ここで、誤差項uは互いに独立に正規分布$N(0,\sigma_u^2)$に従うとする

定期収入、賞与にそれぞれ対応する変数をincome, bonusとして、上記の重回帰モデルを統計ソフトウェアによって最小二乗法で推定したところ、次の出力結果が得られた(次スライド)。なお、出力結果の一部を加工している。また、出力結果のInterceptは定数項α_0を表している

この重回帰モデルに対する解析結果の解釈で以下は正しいか？

1. 賞与を一定としたときに、定期収入が1万円増えると消費支出が約0.39万円増える傾向がある。
2. 賞与と定期収入が同時に1万円増えると消費支出が約0.39万円増える傾向がある。

重回帰モデルの出力結果

```
Coefficients:
             Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  14.58851  2.49814     5.840    5.80e-07
income        0.39461  0.08944     4.412    6.54e-05
bonus         0.47247  0.24370     1.939    0.059
---

Residual standard error: 1.898 on 44 degrees of freedom
Multiple R-squared:  0.5371,    Adjusted R-squared:  0.5161
F-statistic: 25.53 on 2 and 44 DF,  p-value: 4.371e-08
```

消費支出$=\alpha_0+\alpha_1\times$**定期収入**$+\alpha_2\times$**賞与**$+u$

$=14.58851+0.39461\times$ 定期収入 $+0.47247\times$ 賞与$+u$

1. 賞与を一定としたときに、定期収入が1万円増えると消費支出が約0.39万円増える傾向がある。
   →◯
2. 賞与と定期収入が同時に1万円増えると消費支出が約0.39万円増える傾向がある。
   →X (0.39+0.47=0.86万円増える)

例題6　テキスト問題8-1

正規分布表orエクセルを使ってください

とうもろこしの中には黄色と白の２色の実が混ざっていついているものがある。これはバイカラーと呼ばれ、優勢の黄色の実の純系品種と劣勢の白い実の純系品種の１代後輩であるために、その子（純系品種からみたら孫）である種子が持つ2個の遺伝子の黄色、白の組み合わせによって色の違いが生じたものである。この場合、最初の品種が純系であれば黄色と白の実の数の比の期待値はメンデルの法則に従って3:1となるが、そうでないと、例えば花粉が別の品種の花から飛んできたものだったりすると、比率の期待値は変化することになる。
今、あるとうもろこしの実を調べたところ、144粒の実のうち黄色の実が119個、白が25であった。この実が純系種の1代交配から作られたという仮説は棄却できるか、危険率を5%と1%にとって検定しなさい。

テキストappendix E解答を参照してください