Projekt 1 – Analiza trendów

Zadaniem pierwszego projektu „Analiza trendów" jest prezentacja wybranej grupy statków powietrznych, samolotów, śmigłowców czy szybowców budowanych i eksploatowanych na przestrzeni ostatnich kilkudziesięciu lat oraz analiza ich parametrów i prognoza na najbliższe 10 lat. Projekt ten, w formie prezentacji, powinien składać się z następujących części:

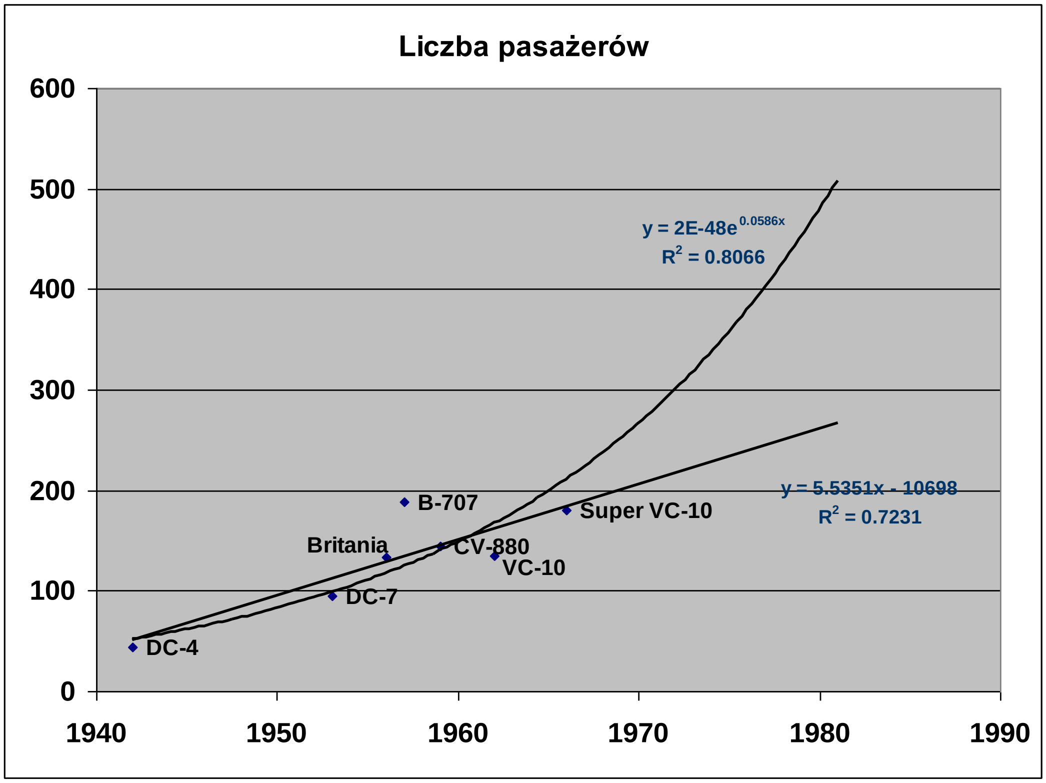
a) definicja wybranej grupy statków powietrznych,
b) prezentacja typowych konstrukcji tej grupy,
c) analiza trendów wybranych danych technicznych i osiągów,
d) podsumowanie, będące próbą określenia istotnych cech samolotu (śmigłowca, szybowca), który byłby produkowany za 10 lat.

Tak przygotowana prezentacja nie powinna trwać dłużej niż 10 minut

Ad. a) - część ta polega na doprecyzowaniu charakterystycznych cech dla wybranej grupy statków powietrznych; np. dla grupy określonej jako „samolot szkolno-treningowy z silnikiem tłokowym", można sprecyzować odbiorcę (użytkownika) samolotu, pokazać typowe wymogi stawiane takim samolotom (np. wymóg startu z lotniska trawiastego), itp.

Ad. b) - prezentacja typowych konstrukcji to część wymagająca uważnej selekcji samolotów z danej grupy; powinny być to konstrukcje, będące typowymi przedstawicielami zdefiniowanej grupy, w danym okresie, produkowane w znaczącej liczbie i stanowiące niemal standard w swojej grupie (np. Embraer Tucano jako przedstawiciel samolotów szkolno-treningowych z silnikiem turbośmigłowym w latach 80-tych). Wskazana jest prezentacja po 2-3 typowe konstrukcje dla okresów 10-letnich. Analiza powinna obejmować 30-40 lat wstecz.

Ad. c) - w tej części należy poddać analizie kilka najważniejszych, dla danej grupy statków powietrznych, danych technicznych (masowych, geometrycznych) oraz parametrów osiągowych; należy je przedstawić w odniesieniu do roku, w którym samolot wszedł do użytku, względnie roku oblotu danego typu. Następnym etapem jest znalezienie krzywych trendu. Można to zrobić wykorzystując stosowne opcje arkuszy kalkulacyjnych czy programów do tworzenia wykresów lub też „ręcznie" stosując regułę minimum kwadratów (patrz poniżej). Jest istotne, aby zbadać jakość krzywej trendu, na przykład wskaźnikiem korelacji liniowej Pearsona (współczynnik R-kwadrat w programie MS Excel). Pamiętać jednak należy, że nie jest to jedyny wskaźnik korelacji a bazując na jednym wskaźniku, przy nie najlepiej dobranych danych możemy wyciągnąć błędne wnioski. Przykład na Rys.1 pokazuje, że mimo lepszego dopasowania krzywej wykładniczej (większe R^2) linia trendu może być pozbawiona sensu - zgodnie z nią liczba pasażerów w typowym samolocie „transatlantyckim" w roku 1980 powinna przekroczyć 500 a 10 lat później osiągnąć niemal 900.



Rys.1 - Przykład analizy trendu z poprawnie (liniowa) i błędnie (wykładnicza) przyjętą funkcją regresji

Ad. d) - na podstawie wyników analizy trendów, należy dokonać próby prognozy na 10 lat do przodu, w wyniku czego powinny powstać założenia dla nowoprojektowanego samolotu; należy podać najważniejsze parametry masowe, względnie geometryczne oraz zakładane osiągi dla nowej konstrukcji.

Podstawy analizy danych

W niniejszym rozdziale zostaną przedstawione podstawowe zależności z zakresu analizy danych, w tym w szczególności sposób wyznaczania funkcji regresji (linii trendu). Rozważmy gromadzone dane jako zbiór par liczb: (argument x_i, wartość Y_i). Załóżmy, że:

$$Y_i = f(x_i) + \varepsilon_i \qquad (1)$$

gdzie:

$f(x_i)$ - nieznana funkcja,

ε_i - błąd losowy o wartości oczekiwanej $E(\varepsilon_i) = 0$.

Przyjmijmy funkcjonał postaci:

$$\mathrm{J}(f) = \sum_{i=1}^{n} (Y_i - f(x_i))^2 \qquad (2)$$

Funkcję, która minimalizuje wartość J(f) nazywamy estymatorem najmniejszych kwadratów nieznanej funkcji regresji f.

Jednym ze wskaźników określających jakość estymacji (przybliżenia) rozkładu doświadczalnego przez funkcję *f* jest współczynnik korelacji liniowej Pearsona, który możemy zapisać następująco:

$$R = \frac{\sum_{i=1}^{n}(f(x_i) - \overline{f(x)})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(f(x_i) - \overline{f(x)})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \qquad (3)$$

gdzie:

$$\overline{f(x)} = \frac{1}{n}\sum_{i=1}^{n} f(x_i) \quad , \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad (4)$$

Im większa wartość współczynnika, tym większa jest zależność liniowa między wartościami doświadczalnymi i przybliżonymi. R = 0 oznacza brak liniowej zależności, R = 1 oznacza dokładną dodatnią liniową zależność. Wielkość „R-kwadrat" (w programie MS Excel) to w istocie kwadrat zdefiniowanego powyżej współczynnika.

Regresja liniowa

Jeżeli przyjmiemy, że nieznana funkcja *f* ma postać funkcji opisującej linię prostą:

$$f(x) = ax + b \qquad (5)$$

to współczynniki tej funkcji spełniające warunek minimum funkcjonału zdefiniowanego we wzorze (2) można wyznaczyć z zależności:

$$a = \frac{n\sum_{i=1}^{n} x_i Y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (6)$$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i Y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (7)$$

Odnosząc powyższe wzory do przykładu pokazanego na Rys.1 zmienna x to czas (rok) a Y to wartość badanego parametru (liczba pasażerów).

Literatura:

1. Siegmund Brandt, Analiza danych - metody statystyczne i obliczeniowe, Wydawnictwo Naukowe PWN, Warszawa 1998
2. Lesław Gajek, Marek Kałuszka, Wnioskowanie statystyczne - modele i metody, Wydawnictwa Naukowo Techniczne WNT, Warszawa 1996
3. Daniel P. Raymer: Aircraft Design: A Conceptual Approach, AIAA Education Series 2012
4. John D. Anderson, Jr.: Introduction to Flight, McGraw Hill, 2005