**MAKALAH SEJARAH STATISTIKA DAN FILSAFAT ILMU**

**“PEARSON CHI-SQUARE AND GOODNESS OF FIT”**



Disusun oleh :

Kelompok 8

1. Awwalina Ghaida R. (13820)
2. Revi Firmansyah (13878)
3. Ratna Mutia K. (14595)
4. Fella Shufa Nur Masfufa (14642)
5. Tri Mutoharoh (14759)
6. Ferri Is Handoko (14784)

**PRODI STATISTIKA JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS GADJAH MADA
YOGYAKARTA
2013**

**PEARSON CHI SQUARE AND GOODNESS OF FIT**

* **BACKGROUND**

**Karl Pearson**(27 Maret 1857 – 27 April  1936) adalah kontributor utama perkembangan awal statistika  hingga sebagai disiplin ilmu tersendiri. Ia mendirikan Departemen Statistika Terapan di University College London pada tahun 1911 yang mana merupakan jurusan statistika pertama kali untuk tingkat universitas di dunia.

Ketertarikan Pearson dalam statistik analitis dimulai pada akhir tahun 1880-an setelah ia menjadi profesor matematika terapan dan mekanik di University College. Karyanya yang terkenal adalah *Pearson chi square dan goodness of fit.*

Awalnya Pearson mengadaptasi matematika mekanika, dengan menggunakan metode momen, untuk membangun sistem statistik baru untuk menginterpretasikan data Weldon yang menghasilkan kurva asimetris, karena tidak ada sistem seperti itu ada pada saat itu. Istilah 'momen' berasal dari mekanik dan merupakan kekuatan sekitar titik rotasi, sementara momen dalam statistik adalah rata-rata. Dari metode momen, Pearson mendirikan empat parameter untuk kurva untuk menunjukkan bagaimana data mengumpul (mean), menyebar (deviasi standar), ketidaksimetrisan (skewness) dan jika bentuk distribusi yang memuncak (kurtosis). Metode momen, yang memungkinkan Pearson untuk menciptakan metodologi statistiknya dianggap penting untuk menafsirkan setiap data statistik, apapun bentuk distribusinya. Oleh karena itu, sistem ini memungkinkan Pearson untuk menganalisis data yang mengakibatkan berbagai bentuk distribusi, dan memungkinkan untuk bergerak melampaui batas kurva normal.

Pearson memutuskan bahwa metode objektif untuk mengukur *goodness of fit* adalah sesuatu yang diinginkan untuk distribusi yang tidak sesuai dengan kurva normal. Pertimbangan Pearson menentukan ukuran uji kebaikan muncul dari kuliahnya pada tanggal 21 November tahun 1893, ketika ia meminta murid-muridnya, 'Bisakah kurva normal selalu cocok dengan suatu kumpulan data? " Jawabannya adalah 'tidak selalu' karena ada banyak jenis data yang dapat menghasilkan kurva asimetris. Pearson pun merancang uji goodness of fit untuk distribusi asimetris menggunakan momen keenam dari metode momen di 1892 .

Quetelet, ilmuwan Belgia, membuat salah satu upaya untuk menyesuaikan satu set data pengamatan ke kurva normal pada tahun 1840. Wilhelm Lexis merancang *Lexican Rasio L* sebagai uji *goodness fit* untuk menentukan apakah distribusi empiris sesuai dengan distribusi normal, sementara Francis Ysidro Edgeworth memberikan kebaikan uji fit tahun 1887 yang didasarkan pada pendekatan normal terhadap distribusi binomial. Meskipun banyak ilmuwan abad kesembilan belas lainnya yang berusaha untuk menemukan uji *goodness of fit*, seperti statistikawan Amerika, Erasmus Lyman de Forest, dan matematikawan Italia, Luigi Perozzo, mereka tidak memberikan dasar teoritis yang mendasari formula mereka, berbeda dengan yang dilakukan Pearson.

Namun dari berbagai ilmuwan yang menemukan goodness of fit, terdapat tentangan melalui Adolphe Quetelet, "bapak statistik," yang percaya bahwa hampir semua fenomena dapat dijelaskan oleh distribusi normal (disediakan hanya jika jumlah sampel yang diperiksa cukup besar), Pearson menurunkan sistem frekuensi kurva umum untuk kepentingan distribusi asimetris. (Dalam hubungan itu, ia mencatat bahwa variabilitas di antara individu-individu adalah konsep yang sangat berbeda dari variasi yang dilakukan untuk mengestimasi nilai tunggal. Dia memperkenalkan istilah deviasi standar dan disimbolkan σ untuk menandakan probabilitas kesalahan.) Setelah menetapkan jenis kurva yang cocok untuk semua jenis pengamatan, Pearson mencari kriteria yang dapat mengukur *goodness of fit*. Sehingga, dalam sebuah makalah terkenal yang diterbitkan pada tahun 1900, ia memperkenalkan *chi square*.

Quetelet menyangkal tentang adanya hubungan antara frekuensi dalam distribusi di sebuah data pengamatan dengan frekuensi yang dihitung berdasarkan asumsi distribusi normal. Mereka tidak memiliki pengukuran atas ketidaksesuaian dari data pengamatan dan harapan. Dengan memperkenalkan chi square, Pearson menyediakan pengukuran di luar distribusi. Chi square ternyata menjadi statistik sangat berguna, dan sekarang menempati posisi utama dalam teori statistik. Namun, tanpa sepengetahuan Pearson, seseorang dari Jerman - Friedrich Helmert - telah menemukan distribusi chi-kuadrat pada tahun 1875, ketika mempelajari distribusi sampling dari variansi sampel, di mana sampel berasal dari populasi normal. Akhirnya, Pearson membuat konteks yang berbeda yaitu, masalah *goodness of fit* yang kemudian diperluas penerapannya pada analisis frekuensi dalam tabel kontingensi. Pearson sendiri tidak jelas bagaimana jumlah yang tepat untuk derajat kebebasan yang digunakan. Dia selalu menggunakan jumlah sampel dikurangi 1 untuk derajat kebebasan. Seperti yang kemudian ditunjukkan oleh Ronald A. Fisher bahwa hasil yang lebih akurat dicapai bila jumlah sampel dikurangi dengan 1 untuk setiap estimasi parameter.

Memang, uji goodness of fit dan chi-square yang dikontribusikan Pearson mewakili teori paling penting untuk teori modern statistik, karena mengangkat substansial praktek statistik matematika. Yang utama pentingnya, distribusi chi-kuadrat dan uji goodness of fit berarti bahwa statistik bisa menggunakan metode yang tidak bergantung pada distribusi normal untuk menginterpretasikan data.

* **BIOGRAFI**

 Karl Pearson lahir pada tanggal 27 Maret 1857 di London, Inggris. Ibunya yang bernama Fanny Smith dan ayahnya William Pearson berasal dari keluarga Yorkshire. Ayahnya adalah seorang pengacara yang sukses.

            Karl Pearson bersama dengan kakak dan adik perempuannya dibesarkan di sebuah keluarga kelas menengah atas. Ayahnya mendorongnya agar dia menumpukan pendidikannya di bidang hukum. Dorongan inilah yang membawa Pearson menekuni bidang hukum di University College School, London.

Setelah menamatkan pendidikannya di bidang hukum, barulah Pearson mulai menjejaki displin matematiknya di King’s College, Cambridge pada tahun 1876. Pearson lulus dari Universitas Cambridge tahun 1879 dan berkat kecemerlangannya dia dianugerahi oleh King’s College, Cambridge sebuah beasiswa di antara 1880 hingga 1886.

Belum lagi merasa puas, dia akhirnya merantau ke Jerman dengan menggunakan beasiswa yang diperolehnya. Selama hampir dua tahun di Jerman, atau lebih tepat lagi di Universitas Berlin dan Heidelberg, Pearson menelaah beberapa displin ilmu seperti fisika, metafisika, falsafah, biologi, sastra klasik Jerman abad ke-16 dan Renaissance, serta tentang darwinisme. Di Universitas Heidelberg dia mempelajari tentang fisika di bawah G. H. Quincke dan metafisika di bawah Kuno Fischer. Sementara di Universitas Berlin dia mempelajari hukum Roma, falsafah, darwinisme, sastra klasik Jerman dan ilmu lainnya. Kegigihan beliau dalam menjelajahi bidang ilmu ini akhirnya berbalas dengan kesuksesannya menguasai bidang-bidang yang ditelaahnya itu. Setelah itu, barulah ia kembali ke Inggris.

Tahun 1880 Pearson menerbitkan buku pertamanya yang berjudul “*The New Werther*” yang berisi tentang alasan mengapa ia mempelajari berbagai macam disiplin ilmu. Pada tahun 1882-1884 ia mengajar di London pada berbagai topik seperti kehidupan sosial Jerman, pengaruh Martin Luther, dan topik sejarah. Ia juga menulis esai, artikel dan ulasan serta menggantikan profesor matematika di King College dan University College London. Sebagian besar karir Pearson dihabiskan di University College London setelah ia diangkat menjadi ketua jurusan matematika terapan tahun 1885 sebagai Profesor Goldsmid.

Pada tahun 1885 Pearson menjadi anggota sebuah perkumpulan pria dan wanita yang didedikasikan untuk diskusi tentang peran pria dan wanita dalam masyarakat. Di perkumpulan itulah dia bertemu dengan Maria Sharpe, wanita yang akhirnya dinikahi Pearson pada tahun 1890. Pernikahan Pearson dengan Maria dikaruniai dua orang putri dan seorang putra, yakni Sigrid Loetitia Pearson, Helga Sharpe Pearson, dan Egon Sharpe Pearson. Anak laki-laki Pearson, Egon Pearson, menjadi ahli statistik seperti ayahnya dan mendirikan Pearson lemma Neyman.

 Ketertarikannya terhadap metode statistika analitik berkembang melalui Weldon and Galton. Pada tahun 1892, Pearson menerbitkan bukunya yang berjudul *“Grammar of Science”* menggambarkan keyakinan yang tumbuh bahwa statistik analitik terletak di semua dasar pengetahuan, karena hampir semua fenomena dapat dijelaskan oleh distribusi normal. Dia juga memperkenalkan istilah standar deviasi dan simbolα untuk yang pertama dan terakhir, dan dalam makalah yang terkenal diterbitkan pada tahun 1900, dia memperkenalkan chi square. Chi square ternyata menjadi statistik yang sangat berguna, dan sekarang menempati posisi utama dalam teori statistik.

Pada tahun 1900, Pearson menjadi salah satu pendiri Biometrika bersama Galton dan Weldon. Sebuah jurnal dikhususkan untuk mempelajari statistika dari masalah Biologis. Dengan ilmu itu ia diangkat sebagai ketua Laboraturium Eugenics. Kemampuan analisis statistiknya berhasil menunjukkan tentang isu-isu penyebab TBC, yakni menunjukkan bahwa TBC lebih terkait ke keturunan daripada factor lingkungan. Pearson mengedit jurnal sampai dia meninggal dan berhasil masuk ke media terkemuka dunia untuk membahas statistika teori dan praktek. Pearson menerima gelar kehormatan dari University of St Andrews dan University of London. Dia juga terpilih sebagai Royal Society of Edinburgh.

* **Warisan, Manfaat, dan Teladan Karl Pearson**

**Warisan**

Selama tahun 1884-1890 Karl Pearson sangat produktif. Selain memberikan pengajaran tentang mekanika, statis dan dinamis, dia melengkapi volume pertama Clifford yang belum selesai, “*The Common Sense of The Exact Sciences*” (diterbitkan tahun 1885), melengkapi dan mengedit sebagian volume pertama Todhunter, “*History of The Theory of Elasticity*” yang kemudian diterbitkan di beberapa paper dalam matematika terapan. Pearson juga mengajar *The Ethic of Free Thought,* melakukan penelitian di berbagai topik sejarah seperti evolusi umat Kristen di barat. Mungkin kita akan berhenti menyadari bahwa Pearson sebagai pendiri statistik karena professor matematika terapan berusia 33 tahun, memiliki reputasi tinggi di berbagai daerah namun masih belum mulai bekerja masalah statistik. Sampai akhirnya ada dua hal yang mengubah karirnya yakni publikasi oleh Galton dari buku *Natural Inheritance* pada tahun 1889 dan penunjukan Weldon sebagai Profesor Zoologi di Universitas College London pada tahun 1890.

Melalui Weldon dan buku Galton, Pearson menjadi tertarik untuk mengembangkan metode matematika untuk mempelajari evolusi dan keturunan. Dari tahun 1893 sampai tahun 1912 Pearson menulis 18 paper berjudul *Mathematical Contributions to the Theory of Evolution*. Paper tersebut berisi tentang analisis regresi, koefisien korelasi, uji chi square (1900). Chi square inilah yang menghapuskan distribusi normal dari titik utama.

Pada tahun 1892 Pearson menulis buku “*The Grammar of Science*” yang menyatakan bahwa analitik statistik merupakan dasar semua ilmu pengetahuan. Dia berusaha memperluas ilmu pengetahuannya ke semua aspek. Pearson menciptakan istilah *standar deviasi* pada tahun 1893. Karyanya dipengaruhi oleh Edgeworth. Sementara beberapa waktu kemudian dia mempengaruhi karya Yule tentang korelasi.

Tidak hanya itu, Pearson bersama dengan Weldon dan Galton menulis jurnal statistika *Biometrika*. Karya tersebut dipresentasikan di Royal Society dan mereka menerima penghargaan Fellow di tahun 1896 dan menerima medali Darwin tahun 1898. Jurnal tersebut berisi tentang kesimpulan biologis yang dianalisis dengan analisis matematik.

Tahun 1906 Pearson berupaya untuk bisa mendirikan pusat pascasarjana. Warisan Pearson terlihat menonjol saat di tahun 1911 dia mendirikan Departemen Statistika Terapan di Universitas College London dan menjadikannya sebagai jurusan statistika pertama kali di dunia di tingkat perguruan tinggi dengan memasukkan Laboratorium Biometrika yang selama dua tahun telah dipegangnya.

Tidak hanya jurnal Biometrika, Pearson juga menerbitkan jurnal *Annals of Eugenics* (sekarang *Annals of Human Genetics*) pada tahun 1925. Dia juga menerbitkan *Drapers Company Research Memoirs* yang sebagian besar memberikan catatan output dari Departemen Statistik Terapan dan tidak dipublikasikan di tempat lain. Sepanjang hayatnya, Pearson telah menulis 18 kertas kerja di bawah tajuk umum, Sumbangan Matematik Pada Teori Evolusi (*Mathematical Contribution to the Theory of Evolution*).

**Manfaat**

Karl Pearson mengembangkan matematika statistika secara luas yang mencakup bidang  biologi, epidemiologi, antropometri, kedokteran dan sosial sejarah. Kontribusinya dalam bidang statistik banyak menopang metode statistik klasik yang umum digunakan sekarang ini diantaranya adalah koefisien korelasi Pearson. Koefisien korelasi Pearson dapat digunakan untuk menyatakan besar hubungan linier antara dua variabel ketika data kuantitatif (data berskala interval atau rasio) dan kedua variabel adalah bivariat yang berdistribusi normal. Selanjutnya kontribusi pearson adalah menemukan korelasi yang sering digunakan oleh peneliti (terutama peneliti yang mempunyai data-data interval  dan rasio) yaitu korelasi Pearson atau *Product Moment Correlation*, yaitu untuk mengetahui derajat hubungan dan kontribusi variabel bebas (independen) dengan variabel terikat (dependen), untuk menyatakan ada atau tidaknya hubungan antara variabel X dengan variabel Y, dan untuk menyatakan besarnya pengaruh variabel satu terhadap yang lainnya yang dinyatakan dalam persen.

* Koefisien korelasi



Kontribusi Pearson yang paling dikenal adalah Pearson uji chi square (1900) yang sering disebut juga dengan *pearson chi square* yang bisa digunakan untuk uji kecocokan (*goodness of fit test*). Uji Chi Square berguna untuk menguji hubungan atau pengaruh dua buah variabel nominal dan mengukur  kuatnya hubungan antara variabel yang satu dengan variabel nominal lainnya, uji kebebasan (*test for independence*), pengujian homogenitas serta pengujian variansi dan standar deviasi populasi tunggal.Selain itu penemuan Pearson yang lain adalah *Principal component analysis (PCA)* yang merupakan prosedur matematika yang menggunakan transformasi orthogonal untuk mengkonversi serangkaian pengamatan variabel yang mungkin berkorelasi menjadi seperangkat nilai-nilai variabel linear berkorelasi. PCA juga merupakan sebuah metode statistik yang banyak digunakan dalam sistem pengenalan pola (wajah, karakter, dan lain-lain) serta kompresi citra.

* Chi square distribution



Teori-teori Pearson yang lain diantaranya adalah sistem pearson pada kurva kontinu, teori sumbangan pada pekali kontigen (*theoretical contributions to the coefficient*), pekali biserial dan tetrakorik (*tectrachoric and biserial coefficients*), ogif, skedastiviti (*scedasticity*), kurtosis, ragam korelasi dan regresi (*multiple correlation and regression*), ratio parsial (*partial ratio*), simpangan baku (*standard deviation*), Chi Distance, p-value (Dalam perkembangannya, banyak peneliti yang sering menggunakan P-value untuk  kriteria ujinya. P-valuelebih disukai dibandingkan kriteria uji lain seperti tabel distribusi dan selang kepercayaan. Hal ini disebabkan karena p-valuememberikan 2 informasi sekaligus, yaitu di samping petunjuk apakah H0 pantas ditolak, p-valuejuga memberikan informasi mengenai peluang terjadinya kejadian yang disebutkan di dalam H0 (dengan asumsi H0 dianggap benar. Definisi p-valueadalah tingkat keberartian terkecil sehingga nilai suatu uji statistik yang sedang diamati masih berarti.), teori tes hipotesis, teori statistik keputusan dan lain-lain.

 **Teladan**

 Pearson disebut sebagai “*father statistic*” dan ilmuwan serba bisa. Kegigihannya dalam menjelajahi berbagai bidang ilmu menjadi teladan untuk kita agar tidak menyerah dalam menguasai satu atau beberapa bidang ilmu. Karya-karyanya yang banyak dan luar biasa hingga berguna di masa sekarang juga merupakan kelebihan dari seorang Pearson. Ide-idenya yang dapat terapresiasi dapat dijadikan contoh untuk kita agar berani menuangkan ide-ide untuk menjadi suatu karya besar, seperti gagasan Pearson yang ingin mendirikan Departemen Statistika Terapan menjadi sebuah jurusan statistika pertama kali di dunia di tingkat perguruan tinggi.

* **LAMPIRAN GAMBAR**

* **KESIMPULAN**

Karl Pearson (27 Maret 1857-27 April 1936)adalah kontributor utama perkembangan awal statistika  hingga sebagai disiplin ilmu tersendiri. Ia mendirikan Departemen Statistika Terapan di University College London pada tahun 1911; yang mana merupakan jurusan statistika pertama kali untuk tingkat universitas di dunia. Pearson menikah dengan Maria Sharpe pada tahun 1890, dan membuahkan 3 anak.

Hasil karya Pearson adalah semua mencakup dalam hal aplikasi yang luas dalam pengembangan statistik matematis, yang mencakup bidang biologi, epidemiologi, antropometri, obat-obatan dan sejarah sosial. Pada tahun 1901, dengan Weldon dan Galton, ia mendirikan jurnal *Biometrika* dimana objeknya adalah mengembangan teori statistik. Dia menjadi editor jurnal ini sampai kematiannya. Dia juga mendirikan jurnal Annals of Eugenics (sekarang *Annals of Human Genetics*) pada tahun 1925. Dia menerbitkan *Drapers Company Research Memoirs* sebagian besar untuk memberikan catatan output dari Departemen Statistik terapan dan tidak dipublikasikan di tempat lain. Selain menerbitkan jurnal-jurnal tentang statistika, dia juga menerbitkan jurnal tentang bidang ilmu lain seperti sosial dan evolusi. Buah pikiran Pearson banyak menopang metode statistik klasik yang umum digunakan sekarang ini. Contoh kontribusinya adalah:

1. Koefisien korelasi
2. Metode momen
3. Sistem pearson pada kurva kontinu
4. *Chi-distance*
5. *P-value*
6. Teori tes hipotesis dan teori statistik keputusan
7. *Pearson chi-square test and goodness of fit*
8. *Principal componen analysis*
* **DAFTAR PUSAKA**

*Eisenhart, C (1974). Karl Pearson. In C. C. Gillispie (Ed.), Dictionary of Scientific Biography, 447-473, New York: Charle Scribners and Sons.*

*Lenin, V.I. (1939). Selected works, vol 11, London: Lawrence Wishart.*

*Lord, C.D. (1995). Karl Pearson, English statistician. In E.J. McMurray (Ed.) Notable twentieth century scientists, 1556-15558, New York, Gale Research, Inc.*

*Pearson, E.S. (1938). Karl Pearson: An appreciation of same aspects of his life and work. Cambridge University Press.*

*Pearson, K. (1892). The grammar of science. London: J. M. Dent and Company.*

*Pearson, K (1897). The chance of death and other studies in evolution, dua jilid, New York: Edward Arnold.*

*Pearson, K. (1901). The ethics of freethought and other addresses and essays. London: Adam and Charles Black.*

*Pearson, K. (1914-1930). The Life, letters and labours of Francis Galton, empat jilid, London: Cambridge University Press.*

*Walker, H. (1968). Karl Pearson. In W. H. Kruskal and J.M Tanur (Eds), 691-698, International Encyclopedia of Statistics. New York: Collier and Macmillan.*

<http://brianstat.blogspot.com/2010/12/karl-pearson.html>

<http://dunia-statistika.blogspot.com/2013/04/tokoh-tokoh-statistika-dunia.html>

<http://www.ilerning.com/index.php?option=com_content&view=article&id=694:korelasi-pearson-product-moment-edit-mar&catid=39:hipotesis&Itemid=70>

[http://www.majalahsains.com/2010/09/karl-pearson-1857-1936-tokoh-matematik-dan intelektual-serba-boleh/](http://www.majalahsains.com/2010/09/karl-pearson-1857-1936-tokoh-matematik-dan%20intelektual-serba-boleh/)

<http://maximaresearch.wordpress.com/2011/12/02/pengambilan-keputusan-dengan-p-value/>

<http://www.swlearning.com/quant/kohler/stat/biographical_sketches/bio14.1.html>

<http://zezenrasyayiyuze.blogspot.com/2012/04/biografi-karl-pearson.html>