

Nama : Al Fitra Nur Ramadhani

NIM : 202210370311264

Mata Kuliah : Data Modeling & Simulation B

Simulasi dan Analisis Distribusi Poisson untuk Pemodelan Jumlah Barang Cacat

Deskripsi

Laporan ini menganalisis penggunaan distribusi Poisson untuk memodelkan jumlah barang cacat dalam proses produksi. Melalui simulasi dengan berbagai nilai parameter, laporan ini mengeksplorasi kesesuaian distribusi Poisson dibandingkan dengan distribusi probabilitas lainnya dalam konteks pengendalian kualitas.

Link GitHub : [DATA-MODELING-AND-SIMULATION/DailyProductionDefectiveGoods at master · alfitranurr/DATA-MODELING-AND-SIMULATION](https://github.com/alfitrannur/DATA-MODELING-AND-SIMULATION/DailyProductionDefectiveGoods)

Penjelasan Studi Kasus dan Variabel Acak

Studi kasus ini memodelkan jumlah barang cacat yang dihasilkan dalam proses produksi per hari. Variabel acak yang dianalisis adalah jumlah barang cacat (X) yang terjadi dalam interval waktu tetap (satu hari produksi). Program melakukan 5 simulasi berbeda dengan parameter lambda (λ) yang bervariasi antara 4,0 hingga 6,0, masing-masing dijalankan untuk periode 30 hari.

Alasan Pemilihan Distribusi Poisson

Distribusi Poisson dipilih karena:

1. Cocok untuk memodelkan jumlah kejadian diskrit dalam interval waktu atau ruang tetap
2. Mengasumsikan kejadian cacat terjadi secara independen
3. Laju kejadian (λ) konstan dalam interval waktu
4. Probabilitas kejadian dalam interval sangat kecil dibandingkan dengan total kemungkinan kejadian

Asumsi ini sesuai dengan karakteristik proses produksi di mana:

- Jumlah cacat dihitung per hari (interval tetap)
- Cacat pada satu barang tidak mempengaruhi cacat pada barang lain
- Peluang cacat relatif kecil dibanding total produksi

Hasil Simulasi dan Analisis per Grafik:

1. **Distribusi Jumlah Barang Cacat per Hari**

Grafik menunjukkan histogram frekuensi relatif jumlah barang cacat untuk kelima simulasi dengan overlay kurva Poisson teoretis. Beberapa observasi:

- Simulasi dengan λ lebih rendah (4,0) cenderung memiliki lebih banyak hari dengan jumlah cacat rendah
- Simulasi dengan λ lebih tinggi (6,0) menunjukkan distribusi yang bergeser ke kanan
- Kurva teoretis (garis putus-putus) secara umum mengikuti pola histogram, meskipun ada beberapa deviasi

Grafik ini mendemonstrasikan bagaimana parameter λ yang berbeda menghasilkan distribusi yang berbeda, sesuai dengan ekspektasi teoretis di mana λ merepresentasikan rata-rata jumlah kejadian.

2. Heatmap Jumlah Barang Cacat per Hari

Heatmap menampilkan jumlah barang cacat untuk setiap hari (baris) dan simulasi (kolom). Dari visualisasi ini:

- Terlihat variasi acak dalam jumlah cacat harian, tanpa pola temporal yang jelas
- Beberapa "hotspot" (warna merah lebih gelap) menunjukkan hari dengan jumlah cacat yang sangat tinggi (hingga 13)
- Simulasi dengan λ lebih tinggi cenderung memiliki lebih banyak sel berwarna gelap (lebih banyak cacat)

Heatmap ini memberikan perspektif dua dimensi tentang distribusi data dan memungkinkan identifikasi outlier atau pola temporal.

3. Perbandingan Distribusi

Boxplot menunjukkan ringkasan statistik distribusi untuk setiap simulasi:

- Median (garis horizontal dalam box) meningkat seiring peningkatan λ
- Interquartile range (tinggi box) bervariasi, dengan simulasi 2 ($\lambda=4,5$) dan simulasi 5 ($\lambda=6,0$) menunjukkan variabilitas yang lebih tinggi
- Whisker menunjukkan range nilai, dengan nilai maksimum mencapai 11-13 untuk simulasi dengan λ lebih tinggi

Visualisasi ini mengonfirmasi bahwa rata-rata dan variasi cenderung meningkat seiring peningkatan nilai λ , konsisten dengan teori distribusi Poisson di mana varians sama dengan mean.

4. Perbandingan Distribusi untuk Data Jumlah Cacat

Grafik ini membandingkan distribusi empiris dari simulasi pertama ($\lambda=4,0$) dengan tiga model distribusi teoretis:

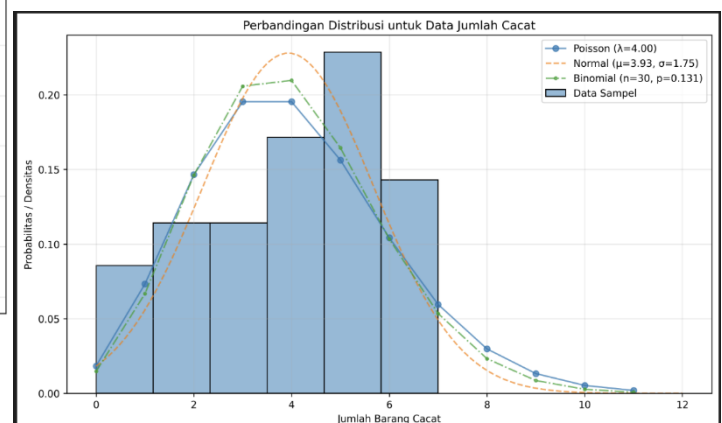
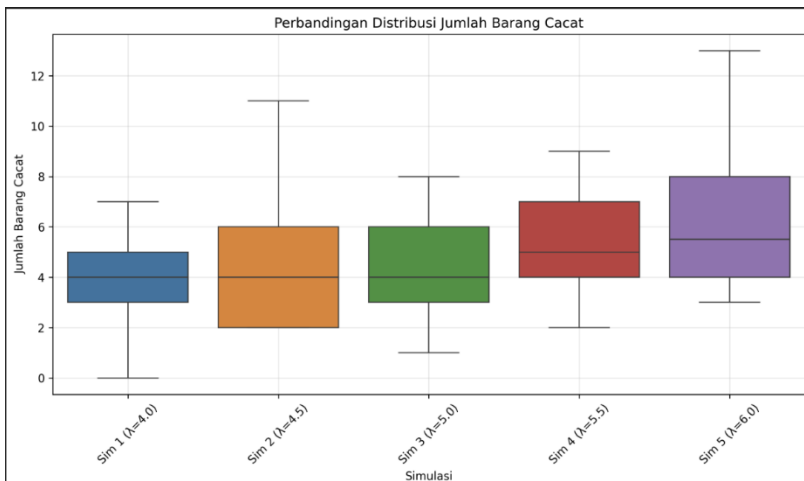
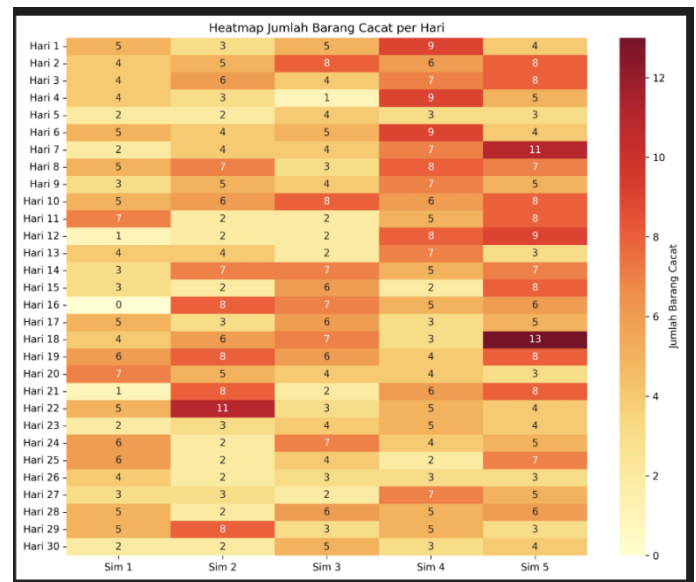
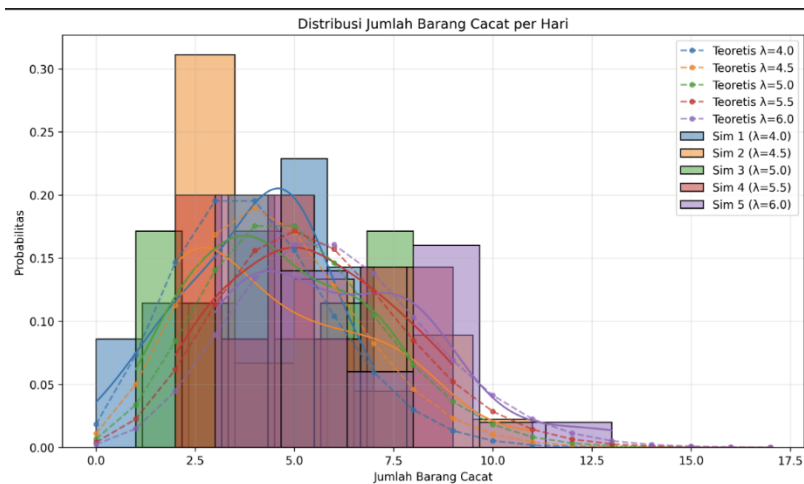
- Poisson ($\lambda=4,0$): Mengikuti data sampel dengan baik, terutama untuk nilai rendah dan tinggi

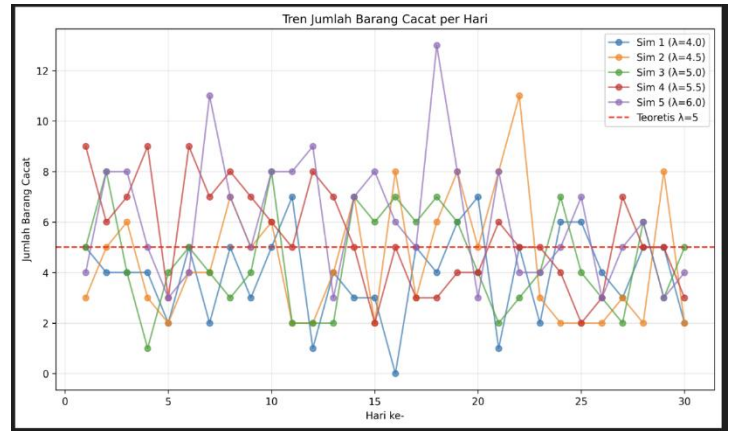
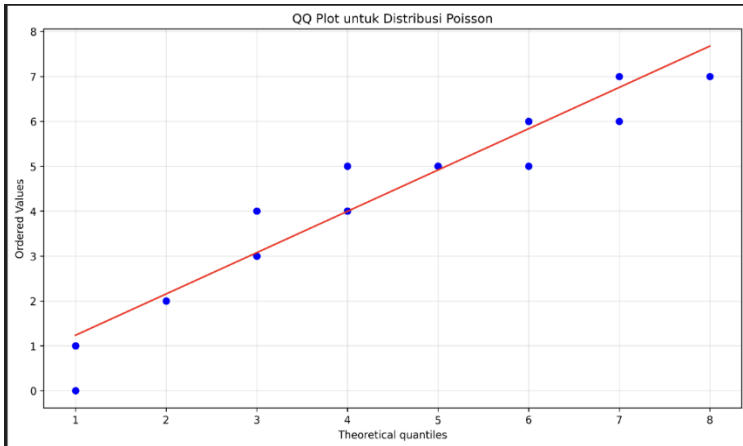
- Normal ($\mu=3,93$, $\sigma=1,75$): Memberikan pendekatan yang cukup baik, terutama di sekitar nilai tengah, tetapi kurang akurat di ekor distribusi
 - Binomial ($n=30$, $p=0,131$): Juga memberikan pendekatan yang serupa dengan Poisson
- Grafik menunjukkan kesesuaian yang baik antara data sampel dan model Poisson, meskipun distribusi normal dan binomial juga menawarkan pendekatan yang dapat diterima.

5. QQ Plot untuk Distribusi Poisson

QQ Plot membandingkan kuantil data empiris dengan kuantil teoretis distribusi Poisson:

- Titik-titik yang mendekati garis merah menunjukkan kesesuaian dengan distribusi Poisson
- Beberapa deviasi terlihat pada kuantil yang lebih tinggi, menunjukkan bahwa data empiris memiliki ekor yang sedikit berbeda dari model teoretis
- Secara keseluruhan, pola titik-titik dekat dengan garis lurus, mengindikasikan kesesuaian yang cukup baik dengan distribusi Poisson





Perbandingan Statistik Empiris vs Teoretis

- **Mean:** Nilai empiris mendekati nilai teoretis untuk semua simulasi
- **Varians:** Beberapa perbedaan terlihat, menunjukkan variasi sampling
- **$P(X=0)$:** Perbedaan signifikan untuk beberapa simulasi karena ukuran sampel terbatas

Pemodelan Sistem Secara Realistik

Distribusi Poisson memodelkan proses cacat dengan baik karena:

- Sesuai dengan asumsi kejadian independen
- Menggunakan interval waktu tetap (satu hari)
- Parameter λ memberikan interpretasi intuitif (rata-rata cacat harian)
- Memungkinkan prediksi probabilitas untuk rentang jumlah cacat

Refleksi Penggunaan Distribusi Lain

- **Normal:** Mungkin baik untuk λ besar, tetapi bisa menghasilkan nilai negatif/non-integer
- **Binomial:** Lebih cocok jika jumlah total item tetap, kurang fleksibel
- **Negative Binomial:** Lebih baik jika data menunjukkan overdispersion (varians > mean), tetapi lebih kompleks