UNIVERSITAS INDONESIA

PENENTUAN DISTRIBUSI DARI BANYAKNYA '*HIT*' KERANDOMAN BARISAN BILANGAN BINER PADA METODE *OVERLAPPING TEMPLATE MATCHING TEST*

SKRIPSI

DHENI TRIADI SUDEWO
0806325503

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
DESEMBER 2011

UNIVERSITAS INDONESIA

PENENTUAN DISTRIBUSI DARI BANYAKNYA '*HIT*' KERANDOMAN BARISAN BILANGAN BINER PADA METODE *OVERLAPPING TEMPLATE MATCHING TEST*

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana sains

DHENI TRIADI SUDEWO
0806325503

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
DESEMBER 2011

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri,

dan semua sumber baik yang dikutip maupun dirujuk

telah saya nyatakan dengan benar.

Nama	: Dheni Triadi Sudewo
NPM	: 0806325503
Tanda Tangan	: 
Tanggal	: Desember 2011

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh

Nama	:	Dheni Triadi Sudewo
NPM	:	0806325503
Program Studi	:	Sarjana Matematika
Judul Skripsi	:	Penentuan Distribusi dari Banyaknya '*Hit*' Kerandoman Barisan Bilangan Biner pada Metode *Overlapping Template Matching Test*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi S1 Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia

DEWAN PENGUJI

Pembimbing	: Sarini Abdullah, M.Stats	(	)
Penguji I	: Dra. Rianti Setiadi, M.Si	(	)
Penguji II	: Dr. Dian Lestari, DEA	(	)
Penguji III	: Dra. Saskya Mary Soemartojo, M.Si	(	)

Ditetapkan di	:	Depok
Tanggal	:	8 Desember 2011

KATA PENGANTAR

Alhamdulillahi rabbil ‘aalamiin, segala puji bagi Allah SWT, Tuhan pencipta alam semesta. Atas ridha dan karunia-Nya lah penulis telah diberikan kesempatan untuk tetap bertahan menghadapi segala rintangan dan cobaan dalam proses kehidupan yang akhirnya menghantarkan penulis menyelesaikan tugas akhir ini.

Penulis menyadari bahwa masih banyak kekurangan yang terdapat dalam tugas akhir ini. Penulis mengharapkan kritik serta saran guna menyempurnakan tugas akhir ini. Pada kesempatan ini, penulis mengucapkan terima kasih kepada orang-orang yang telah sangat berjasa membantu penulis hingga akhirnya tugas akhir ini dapat terselesaikan dengan baik, terutama kepada:

1. Ibu penulis, Waluyo Triningwati dan ayah penulis, Hartoyo Adi Soeprapto, M.Tech yang selalu medoakan serta mendukung penulis.
2. Ibu Sarini Abdullah, M.Stats selaku pembimbing tugas akhir yang telah banyak memberi motivasi penulis, menyediakan waktu, memberi saran serta memberikan banyak ilmu yang sangat berharga dan bermanfaat selama penulis menyelesaikan tugas akhir ini.
3. Ibu Dr. Sri Mardiyati, M.Kom selaku Pembimbing Akademis penulis yang telah mensupport penulis dalam urusan perkuliahan ditiap semesternya maupun saat penulis mulai menjalankan tugas akhir & sidang.
4. Bapak Dr. Yudhi Satria, M.T. selaku Ketua Departemen Matematika, Ibu Rahmi Rusin, S.Si., M.ScTech. selaku Sekretaris Departemen Matematika, Ibu Mila Novita selaku Koordinator Kemahasiswaan, dan Ibu Dr. Dian Lestari selaku Koordinator Pendidikan yang banyak membantu penulis dalam urusan perkuliahan maupun organisasi selama penulis menjalani kuliah disini.
5. Ibu Dra. Rianti Setiadi, M.Si, Dr. Dian Lestari, DEA, dan Dra. Saskya Mary Soemartojo, M.Si selaku penguji kolokium penulis yang bersedia meluangkan waktunya untuk memberi kritik, saran, serta ilmu yang bermanfaat.

6. Seluruh bapak ibu dosen di Departemen Matematika, terutama kepada ibu Dra. Suarsih Utama, Prof. Dr. Belawati HW, Dra. Nora Hariadi, M.Si., Dra. Titin Siswatining, DEA., Dra. Ida Fithriani, M.Si., Dra. Denny Riama Silaban, M.Kom., Dhian Widya, S.Si, M.Kom., Fevi Novkaniza, S.Si, M.Si., Helen Burhan, S.Si., M.Si., Mila Novita, S.Si., M.Si., Dra. Netty Sunandi, M.Si., Dra. Rustina, Dra. Siti Aminah, M.Kom., Dra. Siti Nurrohmah, Dra. Sri Harini, M.Kom., dan kepada bapak Alhaji Akbar, S.Si., M.Sc., Arie Wibowo, S.Si., M.Si., Prof. Dr. Djati Kerami, Hengki Tasman, S.Si., M.Si., Drs. Suryadi Slamet, M.Sc., Drs. Suryadi MT, M.T., Drs. Zuherman Rustam, DEA., dan semua dosen yang tanpa mengurangi rasa hormat tidak dapat disebutkan namanya satu per satu, terima kasih telah mengajar penulis dari tahun pertama hingga tahun akhir serta banyak menyumbangkan ilmu pengetahuan baru yang menarik yang belum pernah penulis dapatkan sebelumnya.
7. Seluruh staff tata usaha serta perpustakaan, Mba Santi, Pak Saliman, Pak Ansori, Mas Salman, Mba Rusmi, Pak Turino, Mas Iwan, yang telah banyak membantu seluruh kegiatan penulis selama disini serta Mas Tatang dan Mas Wawan yang banyak membantu saat penulis membutuhkan bantuan.
8. Kakak-kakak penulis, Fitri Anggraeni Sekardwianti dan Mike Eka Novianti Putri yang selalu memberi semangat kepada penulis.
9. Azki Nuril Ilmiyah, yang selalu menemani serta tempat berbagi keluh kesah baik saat sedih maupun senang. Terima kasih atas perhatian dan bantuannya kepada penulis dalam hal perkuliahan hingga saat penulisan tugas akhir penulis dan terima kasih atas semua yang telah kita jalani bersama.
10. Teman-teman baik penulis Math ‘08, Adhi, Bang Andy, Awe, Arman, Bowo, Arief, Umbu, Maimun, Kiki, Numa, Tute, Ines, Mba Ega, Ade, Dilla, Risya, Sita, Mba Luthfa, Dhea, Aci, Cindy, Ijut, Citra, Nita, dan teman - teman lain yang selalu mendukung Danis, Purwo, Arkies, Agy, Ko Hen, Mas Puput, Dede, Masykur, Juni, Dewe, Mei, Siwi, Vika, Nora, Janu, Resti, Ifah, Eka, Emy, Icha, May TA, Fani, Olin, Yulial & Yulian, Agnes, Maul, Dian, Wulan, Anisah, Uchi D dan Uchi L, terima kasih selama ini telah mengajak penulis

berbagi senyum tawa, pengalaman, cerita, dorongan semangat, maupun berkelana selama penulis disini. *One Math, One Family*!

11. Teman - teman seperjuangan tugas akhir, Mba Luthfa, Dhea, Ijut, Umbu, Fauzan, Kak Sisca (semangat kalian!), Kak Hikma, Kak Zul, Cimz, Bapet, Bang Yos, Andy, Kak Iki, Kak Adi, Kak Siska, Kak Misda, Kak Tika, Kak Siti, dan Kak Putri. Terima kasih atas info-info dan obrolan kita selama menjalani tugas akhir ini.
12. Seluruh kakak angkatan yang bersedia meluangkan waktunya, Kak Ajat (terima kasih banyak atas bantuan pembuktian Proposisi dan Lemmanya), Kak Yanu, Kak Mei, Kak Amri, Kak Angga, Michael 06, Kak Lois, Kak Ar Rizkiyatul, Kak Tino, Kak Winda, terima kasih telah menjadi asdos serta aslab yang memberikan ilmu lebih disamping ibu dan bapak dosen disini.
13. Kakak - kakak angkatan 2007, Adit, Hanif, Anggun, Dhanar, Kak Bowo, Kak Arief, Kak Manda, Kak Dita, Kak Ferdy, Kak Stefi, Kak Nora, Kak Anis, Kak Anjar, Kak Gamar, Kak Isna, dan lainnya. Terima kasih sudah menjadi kakak angkatan yang baik dan banyak memberi bantuan selama ini.
14. Kakak-kakak angkatan 2006, 2005, 2004 dan seluruh penduduk meja putih Hall Math yang selalu berotasi tiap semester, terima kasih telah menjadi teman belajar, berolahraga, organisasi, serta inspirasi bagi penulis dalam kuliah maupun hal lainnya.
15. Adik - adik angkatan 2009, Luthfir, Upi, Budi, Harnoko, Danang, Anton, Agung, Alfian, Dian, Dinda, Michael, Eja & Sofi, Ai, Fitta, Vero, Noe, Sigap, Ica, Ana, Tika, Sondra, Cepi, Soleman, Andrew, Icol, dan lainnya. Terima kasih sudah mengisi hari - hari penulis dengan hal baru selama ada disini. Tetap semangat dan terus berjuang.
16. Adik-adik angkatan 2010, Aid, Pino, Ganesha, Choliq, Mario, Rio, Ihsan, Yudhis, Fariz, Nuel, Barry, Bernard, Marsel, Yandra, Yuza, Dinul, Fikri, Wayan, dan lainnya. Terima kasih sudah menjadi adik angkatan yang ‘asik’ selama disini dan telah membuat penulis menjadi lebih baik melalui acara PDM 2010. Terus berjuang dan tetap semangat.

17. Teman-teman HMD 2010, terutama CT, BPH dan SC, Arman, Andy, Tute, Ines, Ade, Agnes, Aci, Maimun, Maul, Alfian, Mba Ega, Mba Luthfa, Dewe, Kak Tino, dan Kak Arif, terima kasih atas kritik dan saran kepada penulis selama penulis menjalani organisasi maupun perkuliahan sepanjang tahun kedua.
18. Teman-teman, kakak-kakak, dan adik adik dalam seluruh kepanitiaan yang penulis jalani, terutama kepanitiaan MUKER 2009, LOGIKA 2011, SINUS MAPLE, PDM 2010, EKSAKTA, BBM 2008, DISKRET, dan FORSIL, terima kasih telah memberikan penulis wawasan serta pengalaman baru selama menjalani organisasi.
19. Kak Teguh Math 2006, terima kasih atas mentoringnya, berbagi pengalaman, serta kisah-kisah menarik dan tidak monoton yang selalu dibagi tiap minggunya selama kita mentoring.
20. Teman-teman baik penulis di Komplek Kranggan, Fani Rezaniah, Rizky Santika Devi, Fitria Widya Kusuma, terima kasih atas waktu berkumpul dan bergaul bersama.
21. Teman-teman baik penulis SMA a.k.a Panthie, Facur, Dika, Elis, Ajeng, Noni, Khafidz, Gordon, Ardie, Joe, Hendra, Bang Rommy, terima kasih selalu dari SMA hingga saat ini tetap selalu mengajak penulis berkumpul dan berbagi cerita tidak penting bersama.
22. Semua pihak yang telah membantu namun tidak dapat disebutkan satu per satu karena keterbatasan tempat dan daya ingat.

Dan seluruh manusia, teman, keluarga yang penulis kenal, baik saat senang maupun susah. Semoga tugas akhir ini menjadi sesuatu yang bermanfaat bagi siapapun. Maaf atas segala kekurangan yang ada. Terima kasih.

Penulis

2011

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama	: Dheni Triadi Sudewo
NPM	: 0806325503
Program Studi	: Sarjana Matematika
Departemen	: Matematika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis karya	: Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :
Penentuan Distribusi dari Banyaknya '*Hit*' Kerandoman Barisan Bilangan Biner pada Metode *Overlapping Template Matching Test*.

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok
Pada tanggal : 8 Desember 2011
Yang menyatakan

(Dheni Triadi Sudewo)

ABSTRAK

Nama : Dheni Triadi Sudewo
Program Studi : Matematika
Judul : Penentuan Distribusi dari Banyaknya '*Hit*' Kerandoman Barisan Bilangan Biner pada Metode *Overlapping Template Matching Test*

Tugas akhir ini membahas mengenai penentuan distribusi dari banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Mathcing Test*. Metode ini merupakan suatu metode yang terfokus pada sering atau tidaknya muncul 'pola' acak pada tiap blok barisan bilangan biner dengan menggunakan suatu *template*. Penentuan distribusi ini dimulai dengan menggunakan distribusi *Compound Poisson* , lebih khusus lagi menggunakan distribusi *Geometric Poisson.* Lebih lanjut lagi digunakan transformasi *Confluent Hypergeometric Function* (*Kummer's Function*). Selain itu, dalam tugas akhir ini juga diberikan ilustrasi dalam menguji kerandoman barisan bilangan biner dengan menggunakan metode *Overlapping Template Mathcing Test*.

Kata Kunci : distribusi, banyaknya hit, barisan bilangan biner, *Overlapping Template Matching Test*, *Compound Poisson*.
xiv + 57 halaman ; 6 gambar, 7 tabel
Daftar Pustaka : 8 (1972 - 2011)

ABSTRACT

Name : Dheni Triadi Sudewo
Study Program : Mathematics
Title : Detemining Distribution Number of Hit of Bit Sequence Randomness in Overlapping Template Matching Test

This paper discusses about determining distribution number of hit of bit sequence randomness in Overlapping Template Matching Test. This method focusses on how often the pattern appears in each blok of bit sequence by using a template. This determining distribution starts by using *Compound Poisson* distribution, specifically by using Geometric Poisson distribution. Moreover, Confluent Hypergeometric Function is used as transformation's method. Besides, this paper also gives illustration about how to test the randomness of bit sequence using Overlapping Template Matching Test.

Keywords : distribution, number of hit, bit sequence, Overlapping Template Matching Test, *Compound Poisson*.
xiv + 57 pages; 6 pictures, 7 tables
Bibliography : 8 (1972 - 2011)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB 1
PENDAHULUAN

1.1 Latar Belakang

Dalam bidang kriptografi, kerahasiaan suatu informasi merupakan hal yang paling penting. Terdapat banyak metode yang dapat digunakan untuk merahasiakan atau menghasilkan informasi tersebut. Salah satu contohnya adalah metode dalam menghasilkan bilangan acak. Namun, biasanya sulit untuk sebuah program menghasilkan suatu bilangan yang bersifat acak. Maka dari itu untuk menghasilkan suatu bilangan yang bersifat acak, dibutuhkan suatu fungsi/ pembangkit yang dapat membangkitkan bilangan acak tersebut (Otniel, 2011).

Terdapat dua tipe pembangkit yang dapat digunakan untuk menghasilkan bilangan/ barisan bilangan yang bersifat acak: *Random Number Generator* (RNG) dan *Pseudorandom Number Generator* (PRNG). Metode RNG merupakan suatu metode pembangkit barisan bilangan acak dengan sumber nondeterministik. Nondeterministik berarti ketika diberikan nilai awal (*seed*) yang sama, maka pembangkit ini akan menghasilkan barisan *output* yang berbeda - beda. Sedangkan metode PRNG merupakan suatu metode pembangkit barisan bilangan acak dengan sumber deterministik. Deterministik berarti ketika diberikan *input* nilai awal (*seed*) yang sama, maka pembangkit ini akan selalu menghasilkan barisan *output* yang sama (Rukhin, et al., 2001).

Metode RNG merupakan sebuah mekanisme yang digunakan untuk membangkitkan barisan bilangan acak dimana hasil *output*-nya biasanya akan langsung digunakan sebagai *input* pada metode PRNG. Kemudian metode PRNG akan membangkitkan kembali barisan bilangan acak tersebut dengan mekanisme yang sedikit berbeda (Rukhin, et al., 2001).

Kedua metode ini memiliki sedikit perbedaan dalam menjalankan prosesnya. Perbedaan tersebut dapat dilihat pada efisiensi dari kedua metode tersebut. Metode RNG merupakan metode yang kurang efisien jika dibandingkan dengan metode PRNG. Hal ini dikarenakan metode PRNG dapat menghasilkan

barisan bilangan acak dalam waktu yang lebih singkat, deterministik, serta barisan yang dihasilkan bersifat periodik namun tetap acak (Haahr, 1998-2011).

Untuk mengetahui suatu barisan bilangan biner acak atau tidak, perlu dilakukan suatu pengujian. Untuk menguji kerandoman barisan yang dihasilkan oleh metode RNG maupun metode PRNG tersebut terdapat beberapa metode yang dapat digunakan. Salah satu metode tersebut adalah *Overlapping Template Matching Test*.

Pada metode *Overlapping Template Matching Test*, suatu barisan biner dengan panjang n akan dibagi sebanyak N blok dengan panjang masing - masing blok adalah M. Dari masing - masing blok inilah akan diuji apakah ada atau tidak suatu 'pola' yang menjadi acuan (misal untuk $m = 2$ yaitu $11, 10, 01$, atau 00) dengan menggunakan suatu *template* B dengan panjang m. *Template* B merupakan sebuah 'pola' yang digunakan untuk menentukan kerandoman barisan bilangan biner tersebut. Berdasarkan *template* inilah akan diuji kerandoman barisan bilangan biner dengan cara melihat intensitas sering atau tidaknya muncul pola. Apabila terjadi kecocokan (misal *template* yang digunakan 11, dan bagian barisan bilangan yang dievaluasi adalah 11), maka untuk selanjutnya kejadian tersebut disebut '*hit*' (Rukhin, et al., 2001).

Untuk menentukan acak atau tidaknya blok tersebut, akan dilakukan pengujian statistik dengan uji Chi-Square. Pada uji ini barisan bilangan biner tersebut dikatakan acak jika hasil pengamatan intensitas kemunculan bilangan biner pada masing - masing blok mengikuti suatu pola yang diwakili oleh suatu distribusi di bawah asumsi acak. Sehingga pada tugas akhir ini akan ditentukan distribusi dari banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test* apabila asumsi kerandoman terpenuhi.

1.2 Perumusan Masalah dan Ruang Lingkup Masalah

Perumusan masalah dalam tugas akhir ini adalah sebagai berikut:

Bagaimana menentukan distribusi dari banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test*?

Ruang Lingkup dalam tugas akhir ini adalah sebagai berikut:

a) Barisan bilangan biner, pembagian jumlah serta panjang blok, dan *template* pengujian pola acak sudah diberikan.
b) Batas untuk suatu blok mengalami lebih dari N '*hit*' yaitu nilainya akan tetap N '*hit*'.
c) *Template* B yang digunakan khusus untuk $B = 111 \ldots 111$ sepanjang m (*runs of ones*)

1.3 Tujuan Penelitian

Tujuan dari penulisan skripsi ini adalah sebagai berikut:

a) Menjelaskan mengenai penurunan distribusi '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test.*
b) Memberikan ilustrasi dalam menguji kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test*.

BAB 2
LANDASAN TEORI

Pada bab ini akan diberikan dasar - dasar yang digunakan dalam penulisan, yaitu definisi RNG dan PRNG, beberapa distribusi, serta fungsi yang digunakan.

2.1 Definisi *Random Number Generator* (RNG) dan *Pseudorandom Number Generator* (PRNG)

Bilangan acak adalah bilangan yang tidak dapat diprediksi kemunculannya. Pada zaman dahulu, terdapat beberapa cara untuk memperoleh bilangan acak. Diantaranya adalah dengan cara melempar dadu, dan mengocok kartu. Namun pada zaman modern ini (yaitu lebih dari tahun 1940), untuk membentuk bilangan acak dapat dilakukan secara numerik ataupun aritmatik (yaitu dengan menggunakan komputer). Bilangan acak yang dibangkitkan oleh komputer adalah bilangan acak semu (*Pseudo Random Number*) karena menggunakan suatu perumusan tertentu. Karena menggunakan perumusan tersebut, maka bilangan acak yang dihasilkan merupakan bukan merupakan bilangan yang benar - benar acak. Bilangan acak dapat dibangkitkan dengan pola tertentu (dikarenakan tidak ada bilangan acak yang benar - benar acak) dengan mengikuti suatu fungsi distribusi tertentu (Otniel, 2011).

Dalam bidang kriptografi, terdapat dua pembangkit barisan bilangan acak, yaitu *Random Number Generator* (RNG) dan *Pseudorandom Number Generator* (PRNG) dimana keduanya menghasilkan suatu barisan bilangan yang bersifat 'seolah' acak. Dikatakan 'seolah' acak karena tidak ada komputasi yang benar - benar dapat menghasilkan deret bilangan acak secara sempurna. Hal ini disebabkan semua fungsi dalam matematika hanya dapat memetakan satu nilai saja, sedangkan yang diharapkan dari bilangan acak adalah ketika suatu karakter, katakanlah nilai a diacak, akan menghasilkan nilai yang berbeda tiap kali diacak (Otniel, 2011).

2.1.1 *Random Number Generator* (RNG)

Metode *Random Number Generator* (Pembangkit Bilangan Acak), biasa disebut dengan metode RNG, merupakan suatu metode pembangkit barisan bilangan acak dengan sumber nondeterministik. Nondeterministik berarti ketika diberikan nilai awal (*seed*) yang sama, maka pembangkit ini akan menghasilkan barisan *output* yang berbeda - beda. (Rukhin, et al., 2001).

Dalam bidang kriptografi, *output* dari RNG yang digunakan adalah *output* yang tidak dapat diprediksi. Namun dalam kenyataannya, *output* dari RNG terkadang masih dapat diprediksi. Untuk menanggulangi hal ini, yang dapat dilakukan yaitu dengan menggabungkan *output* dari beberapa sumber yang berbeda sebagai *input* dari RNG. Biasanya dibutuhkan waktu yang cukup lama dalam menghasilkan bilangan acak yang memiliki kualitas tinggi.

Bilangan acak dikatakan memiliki kualitas yang baik apabila setelah sekian periode tertentu terjadi perulangan atau munculnya bilangan acak yang sama (semakin lama semakin baik) dan kemunculannya tidak dapat diprediksi. (Rukhin, et al., 2001).

2.1.2 *Pseudorandom Number Generator* (PRNG)

Metode *Pseudorandom Number Generator* (Pembangkit Bilangan Acak Semu), biasa disebut metode PRNG, merupakan suatu metode pembangkit barisan bilangan acak dengan sumber deterministik. Deterministik berarti ketika diberikan *input* nilai awal (*seed*) yang sama, maka pembangkit ini akan selalu menghasilkan barisan *output* yang sama (Rukhin, et al., 2001).

Terdapat hubungan antara kedua metode tersebut, yaitu hasil *output* dari RNG biasanya dapat langsung digunakan sebagai *input* untuk metode PRNG dengan syarat hasil *output*nya sudah memenuhi kriteria acak yang telah diuji dengan uji statistik.

Metode PRNG memiliki beberapa keunggulan apabila dibandingkan dengan metode RNG, yaitu metode ini lebih efisien dan periodik. Lebih efisien

berarti metode ini dapat menghasilkan bilangan acak dalam waktu yang lebih singkat dan periodik berarti secara berkala akan terjadi perulangan atau munculnya bilangan acak yang sama (semakin lama periodenya semakin baik).

Hal tersebut bukan berarti suatu bilangan acak yang mengalami perulangan itu merupakan bilangan acak yang kurang baik. Hal ini dikarenakan setiap bilangan acak akan selalu mengalami perulangan atau munculnya bilangan acak yang sama, namun apabila perulangan tersebut terjadi setelah beberapa periode yang lama, maka bilangan acak tersebut merupakan bilangan acak yang baik. Dalam menghasilkan bilangan acak dengan jumlah besar, lebih baik digunakan metode PRNG. Berikut adalah tabel perbandingan karakteristik kedua metode tersebut:

Tabel 2. 1 Tabel perbandingan karakteristik metode RNG dan PRNG

No	Karakteristik	RNG	PRNG
1	Efisiensi	Kurang Baik	Baik
2	Deterministik	Tidak	Ya
3	Periodik	Tidak	Ya

(Haahr, 1998-2011)

2.2 Macam - macam distribusi

Pada skripsi ini akan digunakan beberapa distribusi dalam proses penurunan distribusi dari banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test*, seperti distribusi Bernoulli, Poisson, Binomial, *Compound Poisson*, dan *Geometric Poisson*.

2.2.1 Distribusi Bernoulli

Percobaan Bernoulli adalah suatu percobaan acak dimana hasil yang mungkin adalah sukses atau gagal. Barisan dari Bernoulli *trials* dikatakan terjadi

apabila percobaan Bernoulli dilakukan berkali - kali dan saling bebas. Lebih lanjut untuk setiap *trial*, probabilitas suksesnya adalah sama yaitu p.

Misalkan variabel acak X yang berhubungan dengan suatu Bernoulli trial, yang didefinisikan sebagai berikut:

$$X(\text{sukses}) = 1 \text{ dan } X(\text{gagal}) = 0.$$

P.d.f dari X dapat ditulis:

$$f(x) = p^x(1-p)^{1-x} \quad , x = 0,1 \qquad (2.1)$$
$$= 0 \quad , \text{yang lainnya.}$$

Maka variabel acak X dikatakan mempunyai *distribusi Bernoulli*.

Ekspektasi dari X:

$$\mu = E(X) = \sum_{x=0}^{1} xp^x(1-p)^{1-x} = (0)(1-p) + (1)(p) = p,$$

dan variansi dari X:

$$\sigma^2 = var(X) = \sum_{x=0}^{1} (x-p)^2 p^x (1-p)^{1-x}$$

$$= p^2(1-p) + (1-p)^2 p = p(1-p).$$

Dengan demikian standar deviasi dari X adalah $\sigma = \sqrt{p(1-p)}$. Distribusi Bernoulli akan dinotasikan dengan $Be(p)$, dengan konstanta p sebagai parameter dari distribusi Bernoulli.

(Hogg & Craig, 1995)

2.2.2 Distribusi Binomial

Variabel acak X dikatakan mempunyai *distribusi Binomial* apabila p.d.f $f(x)$ dari variabel acaknya adalah sebagai berikut:

$$f(x) = \binom{n}{x} p^x(1-p)^{n-x} \quad , x = 0, 1, 2, \dots, n \qquad (2.2)$$
$$= 0 \quad , \text{yang lainnya.}$$

Distribusi binomial akan dinotasikan $b(n, p)$, dengan konstanta n dan p disebut parameter dari distribusi binomial.
(Hogg & Craig, 1995)

2.2.3 Distribusi Poisson

Variabel acak X dikatakan mempunyai *distribusi Poisson* apabila untuk $\lambda > 0$, p.d.f $f(x)$ dari variabel acaknya adalah sebagai berikut:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad , x = 0, 1, 2, \ldots \qquad (2.3)$$
$$= 0 \quad , \text{yang lainnya dimana } \lambda > 0.$$

Distribusi Poisson mempunyai nilai parameter λ dengan nilai $\mu = \sigma^2 = \lambda > 0$. X berdistribusi Poisson dengan parameter λ, dapat ditulis, $X \sim P(\lambda)$.
(Hogg & Craig, 1995)

Berdasarkan Hogg & Craig (1995), distribusi Poisson dengan parameter λ dapat didekati (diaproksimasi) menggunakan distribusi binomial dengan parameter n dan p. Pendekatan ini dapat dilakukan ketika n bernilai besar, dan p sangat kecil sehingga $\lambda = np$.

2.2.4 Distribusi *Compound Poisson*

Berdasarkan Kaas, Govaerts, Dhaene, & Denuit (2002), suatu variabel acak N dikatakan berdistribusi *Compound Poisson*, dengan

$$N = K_1 + K_2 + \cdots + K_M \qquad (2.4)$$

$N = 0$ jika $M = 0$, dimana

- M merupakan variabel acak yang berdistribusi Poisson

- K_m berdistribusi sembarang yang identik dan independen, untuk $m = 1,2,\dots,M$
- M dan K_m independen

Persamaan 2.4 merupakan bentuk dari distribusi *Compound Poisson* yang umum. Dikatakan umum karena distribusi dari variabel acak K_m merupakan sembarang distribusi dengan variabel acak diskrit ataupun kontinu. Kemudian variabel acak M merupakan variabel yang mengukur banyaknya variabel acak K_m yang digunakan untuk menghasilkan variabel acak N. Variabel acak M biasa disebut variabel acak pencampur yang berdistribusi Poisson.

Berdasarkan Nuel (2006), terdapat bentuk khusus dari distribusi *Compound Poisson*, untuk K variabel acak yang diskrit, dimana K menyatakan 'kelas' yang dapat diambil. Variabel acak N dikatakan mempunyai *distribusi Compound Poisson* dengan parameter $(\lambda_k)_{k\epsilon\mathbb{N}^*}$ sedemikian sehingga untuk setiap $\lambda_k > 0$ dan $\sum_{k=1}^{\infty}\lambda_k = \lambda < \infty$ jika

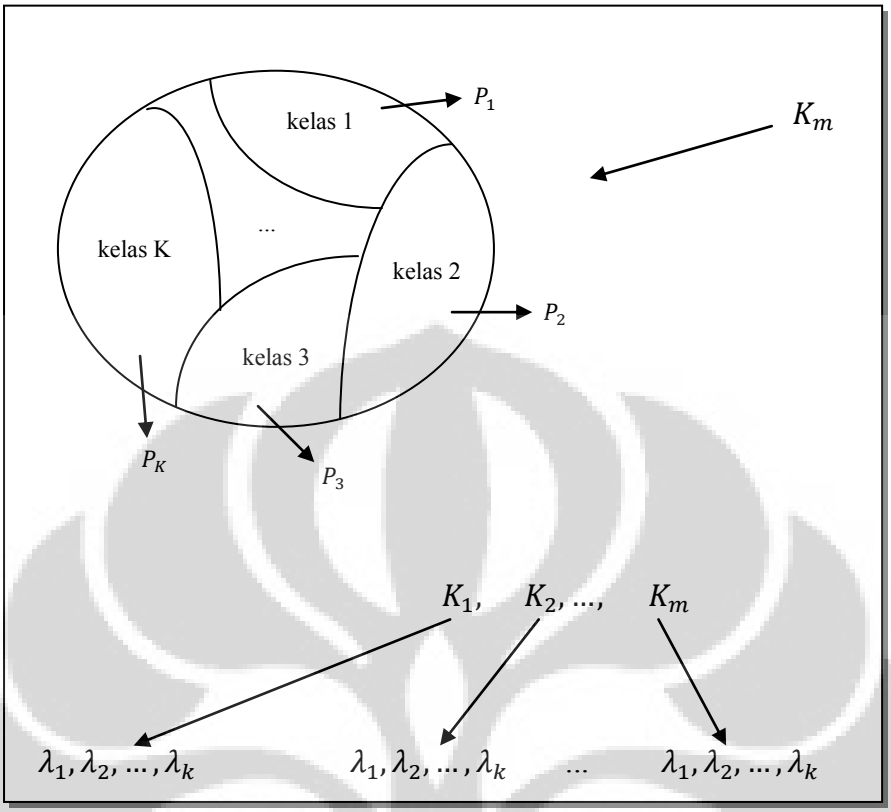
$$N = \sum_{m=1}^{M} K_m = K_1 + K_2 + \cdots + K_M \tag{2.5}$$

dengan

$$P(K = k) = \frac{\lambda_k}{\lambda} \quad \forall k\epsilon\mathbb{N}^* \tag{2.6}$$

dimana

- $M \sim P(\lambda)$ independen terhadap K_m (variabel acak M mengukur banyaknya variabel acak K_m).
- K_i, K_j berdistribusi identik dan independen.
- $\mathbb{N}^*$ didefinisikan sebagai suatu himpunan bagian dari bilangan asli yang khusus.
- λ_k adalah parameter untuk setiap K_m, yang menyatakan 'bobot' bahwa $K_1, K_2, \dots, K_m$ masuk ke kelas K.



Gambar 2. 1 Penggambaran hubungan K_m dengan paramer λ_k

Distribusi *Compound Poisson* dinotasikan $CP((\lambda_k)_{k\epsilon\mathbb{N}^*})$ dengan parameter $(\lambda_k)_{k\epsilon\mathbb{N}^*}$.
(Nuel, 2006)

Lemma 2.1

Jika $N \sim CP((\lambda_k)_{k\epsilon\mathbb{N}^*})$ dengan $\sum_{k=1}^{\infty} \lambda_k = \lambda$ maka $\forall n \in \mathbb{N}^*$

$$P(N = n) = \sum_{m=1}^{n} e^{-\lambda} \frac{\lambda^m}{m!} \sum_{k_1,\ldots,k_m \in \mathbb{N}^*} \mathrm{I}_{\{k_1+k_2+\cdots+k_m=n\}} \frac{\lambda_{k_1} \times \ldots \times \lambda_{k_m}}{\lambda^m} \qquad (2.7)$$

dan $P(N = 0) = e^{-\lambda}$.

(Nuel, 2006)

Penjelasan mengenai perumusan 2.7, terdapat pada bukti di Lampiran 1.

2.2.5 Distribusi *Geometric Poisson* (*Pólya-Aeppli*)

Variabel acak N berdistribusi *Geometric Poisson* (*Pólya-Aeppli*) dengan parameter $\theta \in [0,1]$ untuk bagian geometrik dan parameter $\lambda > 0$ untuk bagian Poisson (dinotasikan $GP(\lambda, \theta)$), jika $N \sim CP((\lambda_k)_{k \in \mathbb{N}^*})$ dengan

$$\lambda_k = \lambda(1-\theta)^{k-1}\theta \quad \forall k \in \mathbb{N}^* \tag{2.8}$$

(Nuel, 2006)

Penjelasan dari definisi *Geometric Poisson* (*Pólya-Aeppli*) di atas yaitu jika terdapat suatu variabel acak N, dimana

1. N berdistribusi *Compound Poisson* dengan parameter λ_k untuk $k \in \mathbb{N}^*$.
2. Terdapat suatu parameter λ dan θ yang masing - masing dapat dikatakan sebagai parameter untuk 'bagian' Poisson dan Geometrik.
3. $\lambda_k = \lambda(1-\theta)^{k-1}\theta \quad \forall k \in \mathbb{N}^*$

Maka variabel acak N tersebut dikatakan berdistribusi *Geometric Poisson* dengan parameter λ_k $k \in \mathbb{N}^*$. Bentuk distribusi *Geometric Poisson* merupakan bentuk khusus dari distribusi *Compound Poisson.*

Berikut dengan menggunakan persamaan 2.7 (bentuk distribusi dari *Compound Poisson*) dan persamaan 2.8 (bentuk dari λ_k pada *Geometric Poisson*), maka akan dihasilkan bentuk distribusi dari *Geometric Poisson* pada Proposisi 2.1 sebagai berikut:

Proposisi 2.1

Jika $N \sim GP(\lambda, \theta)$ maka $\forall\, n \geq 1$

$$P(N=n) = \sum_{m=1}^{n} e^{-\lambda} \frac{\lambda^m}{m!} (1-\theta)^{n-m} \theta^m \binom{n-1}{m-1} \tag{2.9}$$

dengan $\binom{n}{k}$ merupakan koefisien binomial ($_nC_k$), dan $P(N=0) = e^{-\lambda}$.

(Nuel, 2006)

Penjelasan mengenai perumusan 2.9, terdapat pada bukti di Lampiran 2.

Penjelasan Proposisi 2.1 ini yaitu apabila suatu variabel acak berdistribusi *Gemoetric Poisson* dengan parameter λ dan θ, maka dapat dibentuk p.d.f seperti diatas dimana bentuk p.d.f tersebut didapat dengan menurunkan bentuk p.d.f dari suatu variabel acak yang berdistribusi *Compound Poisson* yang ada pada Lemma 2.1. Dimana dengan menggunakan $\lambda_k = \lambda(1-\theta)^{k-1}\theta$,

$$\sum_{k_1,\dots,k_m \in \mathbb{N}^*} \frac{\lambda_{k_1} \times \dots \times \lambda_{k_m}}{\lambda^m} \to (1-\theta)^{n-m}\theta^m$$

dan

$$\sum_{k_1,\dots,k_m \in \mathbb{N}^*} \mathrm{I}_{\{k_1+k_2+\cdots+k_m=n\}} \to \binom{n-1}{m-1}$$

Sehingga didapat bentuk p.d.f dari variabel acak yang berdistribusi *Geometric Poisson* seperti persamaan 2.9. Untuk penjelasan lebih lengkapnya dapat dilihat pada Lampiran 2.

2.3 *Confluent Hypergeometric Function* (*Kummer's Function*)

Bentuk *Confluent Hypergeometric Function* (biasa disebut dengan *Kummer's Function*) berdasarkan Abramovitz & Stegun (1972),adalah:

$$M(a,b,z) = 1 + \frac{az}{b} + \frac{(a)_2 z^2}{(b)_2 2!} + \cdots + \frac{(a)_n z^n}{(b)_n n!} + \cdots \qquad (2.10)$$

dimana

$$(a)_n = a(a+1)(a+2)\dots(a+n-1)$$

$$(a)_0 = 1$$

Notasi lain untuk $M(a,b,z)$ yaitu ${}_1F_1[a;b;z]$ dan $\Phi(a,b,z)$.

Transformasi *Kummer*

$$\Phi(a,b,z) = e^{z}\Phi(b-a,b,-z) \tag{2.11}$$

(Abramowitz & Stegun, 1972)

Setelah didapat bahwa $N \sim GP(\lambda,\theta)$ dengan $\lambda > 0$ dan $\theta \in [0,1]$, maka berdasarkan Abramovitz & Stegun (1972), Proposisi 2.1 dapat ditulis dengan menggunakan *Confluent Hypergeometric Function* seperti yang tercantum pada Proposisi 2.2 berikut:

Proposisi 2.2

$\forall n \in \mathbb{N}^*$ dengan $N \sim GP(\lambda,\theta)$ dengan $\lambda > 0$ dan $\theta \in [0,1]$ didapat

$$P(N=n) = e^{-\lambda}(1-\theta)^{n} z e^{-z}\Phi(n+1,2,z) \tag{2.12}$$

dimana Φ merupakan *Confluent Hypergeometric Function* dan $z = \frac{\lambda\theta}{1-\theta}$.

(Nuel, 2006)

Penjelasan mengenai perumusan 2.12, terdapat pada bukti di Lampiran 3.

Sehingga pada penulisan tugas akhir ini *Confluent Hypergeometric Function* digunakan sebagai salah satu cara untuk merepresentasikan bentuk p.d.f (pada persamaan 2.7) dari distribusi *Compound Poisson.* Selain itu dalam penggunaannya bentuk ini sudah terdapat dalam perangkat lunak yang biasa digunakan. Sehingga untuk keperluan numeriknya hanya diperlukan input-input parameter a, b, z dan akan diperoleh hasil, dibandingkan harus menghitung bentuk $\binom{n-1}{m-1}$ dan $(1-\theta)^{n-m}\theta^{m}$.

BAB 3
PENENTUAN DISTRIBUSI DARI BANYAKNYA '*HIT*' KERANDOMAN BARISAN BILANGAN BINER PADA METODE *OVERLAPPING TEMPLATE MATCHING TEST*

Pada bab 3 ini, akan dibahas mengenai penentuan distribusi dari banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test*.

3.1 Pengujian Kerandoman Barisan Bilangan Biner pada Metode *Overlapping Template Matching Test*

Metode *Overlapping Template Matching Test* ini adalah salah satu metode yang digunakan untuk melihat pola pada barisan bilangan biner. Pola yang dimaksud adalah kemunculan $11 \ldots 111$ sepanjang m pada barisan bilangan biner, untuk selanjutnya akan disebut dengan '*hit*'. Maksudnya yaitu '*hit*' merupakan sebutan untuk barisan bilangan biner yang memiliki kecocokan (kesamaan) dengan *template* acuan yang digunakan. Misal template yang digunakan adalah 111, dengan barisan bilangan biner yang sedang dievaluasi adalah 111, maka karena terjadi kecocokan antara template dengan barisan bilangan yang sedang dievaluasi untuk selanjutnya akan disebut '*hit*'. Berdasarkan pola tersebut, pada akhirnya akan dapat ditentukan apakah suatu barisan biner tersebut random atau tidak.

- Notasi

  ε = barisan bilangan biner dari yang dibangkitkan oleh RNG atau PRNG.

  n = panjang barisan bilangan biner ε.

  M = panjang *bit* tiap blok yang dibentuk dari ε.

N = banyaknya blok yang dibentuk dari ε, dipilih berdasarkan nilai dari M.

B = m-*bit template* (pola) yang digunakan untuk menguji tiap blok.

m = panjang *bit* (bilangan biner) pada *template.*

- Ilustrasi

$$\varepsilon = \underbrace{\varepsilon_1\varepsilon_2 \ldots \overbrace{\varepsilon_M}^{B}}_{\text{Blok 1}}\underbrace{\boldsymbol{\varepsilon_{M+1}\varepsilon_{M+2}\ldots\varepsilon_{2M}}}_{\text{Blok 2}}\varepsilon_{2M}\underbrace{\varepsilon_{2M+1}\ldots\varepsilon_{3M}}_{\text{Blok } N}\ldots\varepsilon_{n-1}\varepsilon_n,$$ dimana

B = *template* dengan panjang bit m, dengan $n = NM$.

Akan didefinisikan v_i yaitu banyaknya blok yang memuat i '*hit*' dari sejumlah blok yang telah dievaluasi. Berikut ilustrasi mengenai v_i:

$$\varepsilon = \underbrace{\mathbf{1011100101}}_{\text{Blok 1}}\underbrace{0110010100}_{\text{Blok 2}}\underbrace{\mathbf{1010110110}}_{\text{Blok 3}}\underbrace{0101010011}_{\text{Blok 4}}$$

$n = 40, M = 10, N = 4.$

Misal digunakan $m = 2$, *template* $B = 11$, maka untuk blok 1:

Tabel 3. 1 Tabel evaluasi blok 1

Posisi *bit*	*Bit*	Akumulasi $B = 11$
$1-2$	10	0
$2-3$	01	0
$3-4$	11	'*hit*' ke -1
$4-5$	11	'*hit*' ke -2
$5-6$	10	0
$6-7$	00	0
$7-8$	01	0

$8-9$	10	0
$9-10$	01	0

Terjadi 2 kali '*hit*' dengan menggunakan $B = 11$, sehingga $v_2 = 1, v_0 = v_1 = v_3 = v_4 = 0$, yang berarti:

- $v_2 = 1$ artinya ada 1 blok (yaitu blok ini sendiri) yang terjadi 2 kali '*hit*' dari blok yang sudah dievaluasi.
- $v_0 = 0$ artinya tidak terjadi '*hit*' dari blok yang sudah dievaluasi.
- $v_1 = v_3 = v_4 = 0$ artinya tidak terjadi 1, 3, 4, dan 5 kali '*hit*' dari blok yang sudah dievaluasi.

Didefinisikan

$$v_i = \sum_{r=1}^{N} v_{i_r}$$

dimana r menyatakan blok pada barisan bilangan biner, dengan $r = 1, 2, \dots, N$.

Apabila terdapat sejumlah N blok, maka akan terdapat sebanyak $N + 1$ untuk nilai v_i, sedemikian sehingga

$$v_{i_r} = \begin{cases} 1, & \text{jika blok } r \text{ mengalami } i \text{ '}hit\text{'}, \ i = 0, 1, 2, \dots, N-1 \\ 0, & \text{untuk yang lainnya} \end{cases}$$

$$v_{N_r} = \begin{cases} 1, & \text{jika blok } r \text{ mengalami '}hit\text{'} \ \geq N \\ 0, & \text{untuk yang lainnya} \end{cases}$$

$$\sum_{i=0}^{N} v_i = N \qquad (3.1)$$

Mengacu pada pendefinisian v_i, dimana v_i menyatakan banyaknya blok yang memuat i '*hit*', maka $\sum_{i=0}^{N} v_i$ menyatakan jumlah blok yang memuat seluruh kemungkinan nilai '*hit*' setelah dievaluasi secara menyeluruh. Sehingga $\sum_{i=0}^{N} v_i$ dapat diyatakan sebagai total blok, yaitu N seperti yang dinyatakan pada persamaan 3.1.

Artinya pada suatu barisan dengan banyak blok adalah N, untuk sejumlah '*hit*' yaitu i, maka akan dievalusi apakah tiap blok pada barisan tersebut

mengalami sejumlah i '*hit*' tersebut. Misalkan sebuah barisan dengan banyak blok 5, akan diuji ada berapa banyak blok yang mengalami 3 '*hit*'. Apabila dari blok 1 hingga blok 5, hanya blok 2 yang tidak mengalami 3 '*hit*' (terdapat 4 blok yang mengalami 3 '*hit*'), maka dituliskan

$$v_3 = \sum_{r=1}^{5} v_{3_r}$$

$$= v_{3_1} + v_{3_2} + v_{3_3} + v_{3_4} + v_{3_5}$$

$$= 1 + 0 + 1 + 1 + 1$$

$$= 4$$

Sehingga didapat $v_3 = 4$, yang berarti terdapat 4 blok yang mengalami 3 '*hit*' setelah mengevaluasi seluruh blok pada barisan.

Berdasarkan cara mengevaluasi yang sama seperti yang dilakukan pada Blok 1 dengan $m = 2$, *template* $B = 11$, didapat:

- Blok $2 = 0110010100$

  Terjadi 1 kali '*hit*' pada blok 2, maka didapat

  $$v_1 = 1, v_2 = 1, v_0 = v_3 = v_4 = 0$$

- Blok $3 = 1010110110$

  Terjadi 2 kali '*hit*' pada blok 3, maka didapat

  $$v_2 = 2, v_1 = 1, v_0 = v_3 = v_4 = 0$$

  sehingga terlihat disini yang awalnya $v_2 = 1$ menjadi $v_2 = 2$ dikarenakan setelah dievaluasi terdapat 2 blok yang terjadi 2 kali '*hit*'.

- Blok $4 = 0101010011$

  Terjadi 1 kali '*hit*' pada blok 4, maka didapat

  $$v_1 = 2, v_2 = 2, v_0 = v_3 = v_4 = 0.$$

  Sama seperti v_2 tadi, nilai $v_1 = 1$ menjadi $v_1 = 2$ dikarenakan setelah dievaluasi terdapat 2 blok yang terjadi 2 kali '*hit*'.

Sehingga setelah semua blok dievaluasi, didapat:

$$v_0 = v_3 = v_4 = 0, v_1 = 2, v_2 = 2$$

Didefinisikan:

$v_i =$ banyaknya blok yang memuat i '*hit*'.

$E_i =$ ekspektasi dari v_i di bawah asumsi barisan tersebut acak.

Dari kedua definisi tersebut, terlihat bahwa keduanya merupakan komponen yang mengevaluasi, yaitu v_i, serta ekspektasi dari v_i, yaitu E_i, barisan berdasarkan banyaknya blok yang memuat i '*hit*'. Namun pada penulisan tugas akhir ini, dalam perumusan masalah hanya disebutkan kerandoman banyaknya i '*hit*' pada barisan bilangan biner. Untuk selanjutnya akan diberikan ilustrasi mengenai perubahan definisi dari yang awalnya berdasarkan banyaknya blok yang mengalami i '*hit*', hanya menjadi banyaknya i '*hit*' saja.

Pertama - tama perhatikan E_i, yaitu ekspektasi dari banyaknya blok yang didalamnya memuat i '*hit*' di bawah asumsi barisan tersebut acak. Dapat dituliskan:

$$E_i = N\pi_i$$

dimana:

$\pi_i =$ probabilitas bahwa suatu blok (diantara blok - blok dalam barisan) akan memuat i '*hit*'

$N =$ banyaknya blok dalam barisan bilangan biner

Kemudian perhatikan v_i, yaitu banyaknya blok yang memuat i '*hit*'. Pada definisi v_i, tidak diketahui blok mana yang sebenarnya didalamnya mengalami i '*hit*.

Demikian pula dengan π_i sebagai probabilitas bahwa suatu blok akan memuat i '*hit*'. Perhatikan bahwa pendefinisian π_i, yang menjadi perhatian adalah banyaknya '*hit*' di dalam suatu blok (v_i) ataupun probabilitas kemunculan i '*hit*'

di dalam suatu blok (π_i) tanpa memperhatikan letak blok. Secara total terdapat N blok dalam barisan bilangan biner, maka $N\pi_i$ menyatakan ekspektasi dari banyaknya blok yang memuat i '*hit*' di bawah asumsi barisan tersebut acak.

Dari penjelasan tersebut, dapat dinyatakan definisi kerandoman yang semula didefinisikan sebagai banyaknya blok yang di dalamnya terjadi i '*hit*', dapat juga dinyatakan sebagai banyaknya '*hit*' saja (atau banyaknya i '*hit*'). Hal ini dikarenakan banyaknya blok yang di dalamnya mengalami i '*hit*' adalah sembarang, baik untuk v_i maupun untuk $E_i = N\pi_i$.

Selanjutnya untuk menguji barisan ε acak atau tidak, akan dilakukan perbandingan antara v_i dan E_i. Berdasarkan kedua definisi ini, apabila v_i mendekati E_i berarti barisan tersebut dikatakan acak. Digunakan hipotesis sebagai berikut:

H_0 : Barisan ε random

Hipotesis di atas dapat dinyatakan dalam suatu perumusan. Jika barisan ε random, maka W, yaitu banyaknya '*hit*' pada suatu blok, adalah berdistribusi *Compound Poisson*. Sehingga H_0: barisan ε random, dapat dituliskan kembali sebagai berikut:

$$P(W = 0) = \pi_0 = e^{-\eta}$$

$$P(W = i) = \pi_i = \frac{e^{-\eta}}{2^i}\sum_{l=1}^{i}\binom{i-1}{l-1}\frac{\eta^l}{l!} = \frac{\eta e^{-2\eta}}{2^u}\Phi(i+1, 2, \eta); \text{untuk } i \geq 1, \eta = \frac{\lambda}{2},$$

H_1 : Tidak demikian

Statistik uji:

$$\chi^2 = \sum_{i=0}^{N}\frac{(O_i - E_i)^2}{E_i}$$

$$= \sum_{i=0}^{N}\frac{(v_i - N\pi_i)^2}{N\pi_i}$$

dimana:

O_i merupakan banyaknya blok yang memuat i '*hit*'.

E_i merupakan ekspektasi banyaknya blok yang memuat i '*hit*'.

N merupakan banyaknya blok yang dibentuk dari ε.

Penentuan distribusi dari π_i inilah yang akan menjadi permasalahan dalam penulisan tugas akhir ini. Berikut adalah prosedur perhitungan untuk nilai π_i:

$$\pi_i = P(W = i)$$

$$= \frac{e^{-\eta}}{2^i}\sum_{l=1}^{i}\binom{i-1}{l-1}\frac{\eta^l}{l!} = \frac{\eta e^{-2\eta}}{2^i}\Phi(i+1,2,\eta); \text{untuk } i \geq 1, \eta = \frac{\lambda}{2}, \quad (3.2)$$

dimana

$l = 1, 2, \ldots i$

$$\lambda = \frac{M - m + 1}{2^m}$$

Pandang bentuk berikut

$$\Phi(a,b,z) = 1 + \frac{az}{b} + \frac{(a)_2 z^2}{(b)_2 2!} + \ldots + \frac{(a)_n z^n}{(b)_n n!} + \ldots, \text{dimana}$$

$$(a)_n = a(a+1)(a+2)\ldots(a+n-1)$$

$(a)_0 = 1.$

$\Phi(a,b,z)$ dikenal dengan *Confluent Hypergeometric Function* (*Kummer's function*) yang telah dibahas pada subbab sebelumnya.

Aturan Keputusan:

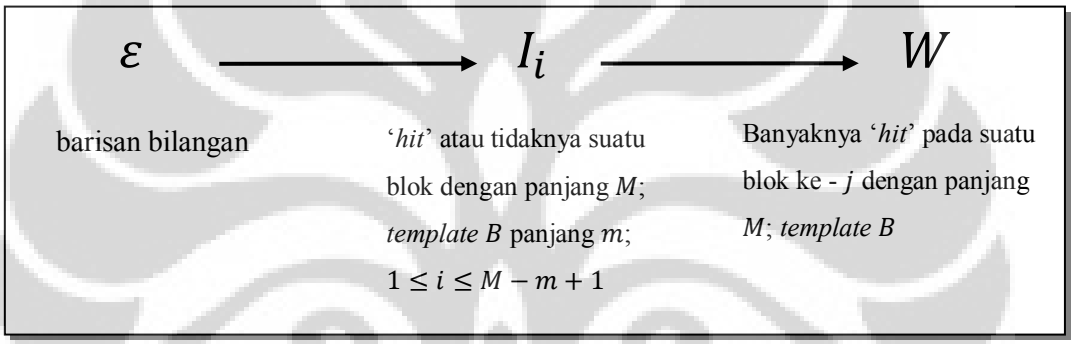
Pada tingkat signifikasi α, H_0 ditolak jika $\chi^2 \geq \chi^2_{((N+1)-1)} \cdot \chi^2_{((N+1)-1)}$ didapat dari tabel Chi-Square, berdasarkan Hogg & Craig (1995) dengan derajat bebas $(N+1) - 1 = N$.

Dengan tingkat signifikasi α, ditolak atau tidaknya hipotesis H_0 yang digunakan bergantung pada nilai $\chi^2_{(N)}$. Dengan melakukan pengujian tersebut, barulah dapat ditentukan apakah barisan bilangan ε random atau tidak.

3.2 Distribusi W

3.2.1 Komponen - komponen dan Pendefinisian W

Akan dijelaskan terlebih dahulu definisi dari W. Gambaran pendefinisian dari W adalah sebagai berikut:



Gambar 3. 1 *Flowchart* pendefinisian W

Pertama - tama diberikan sebuah barisan bilangan biner ε dengan panjang n, yaitu:

$$\varepsilon = \varepsilon_1\varepsilon_2 \dots \varepsilon_M\varepsilon_{M+1} \dots \varepsilon_{2M}\varepsilon_{2M+1} \dots \varepsilon_{2M+1} \dots \varepsilon_{3M} \dots \varepsilon_{n-1}\varepsilon_n \qquad (3.3)$$

dimana

$$\varepsilon_i = \begin{cases} 1 \text{ , dengan probabilitas } p \\ 0 \text{ , dengan probabilitas } q, \text{ dimana } q = 1 - p \end{cases}$$

ε_i adalah percobaan Bernoulli dengan ε_i dikatakan sukses jika $\varepsilon_i = 1$. ε adalah barisan dari Bernoulli *trials* karena merupakan percobaan Bernoulli yang dilakukan berkali-kali dan saling bebas. Lebih lanjut, pada setiap *trial* probabilitas suksesnya adalah sama yaitu p, sehingga dapat dikatakan bahwa:

$$\varepsilon_i \sim \text{Be}(p), 1 \le i \le n$$

Kemudian definisikan I_i untuk blok dengan panjang M dan *template B* dengan panjang m yaitu:

$$I_i = \prod_{j=i}^{i+m-1} \varepsilon_j \,,\ 1 \leq\ i \leq M - m + 1 \tag{3.4}$$

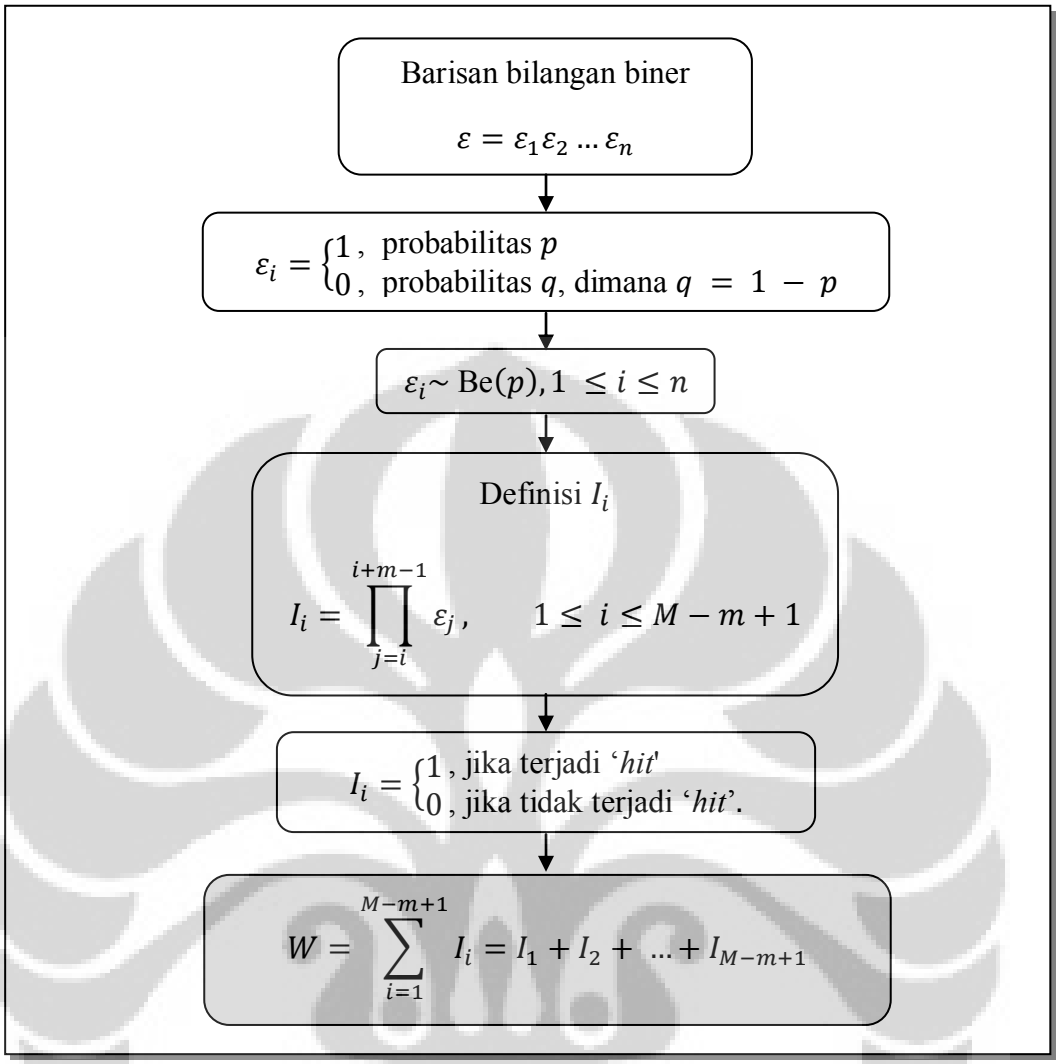
dimana

$I_i = \begin{cases} 1 \text{ , jika terjadi '}hit\text{'} \\ 0 \text{ , jika tidak terjadi '}hit\text{'}. \end{cases}$

Dapat didefinisikan W untuk blok dengan panjang M dan *template B* dengan panjang m pada blok ke - j yaitu:

$$W = \sum_{i=1}^{M-m+1} I_i = I_1 + I_2 + \ \ldots + I_{M-m+1} \tag{3.5}$$

untuk $i = 1, 2, \ldots, M - m + 1$.



Gambar 3. 2 *Flowchart* Komponen *W*

Sehingga didapat komponen - komponen serta pendefinisian dari W. Selanjutnya akan dibahas mengenai distribusi dari W.

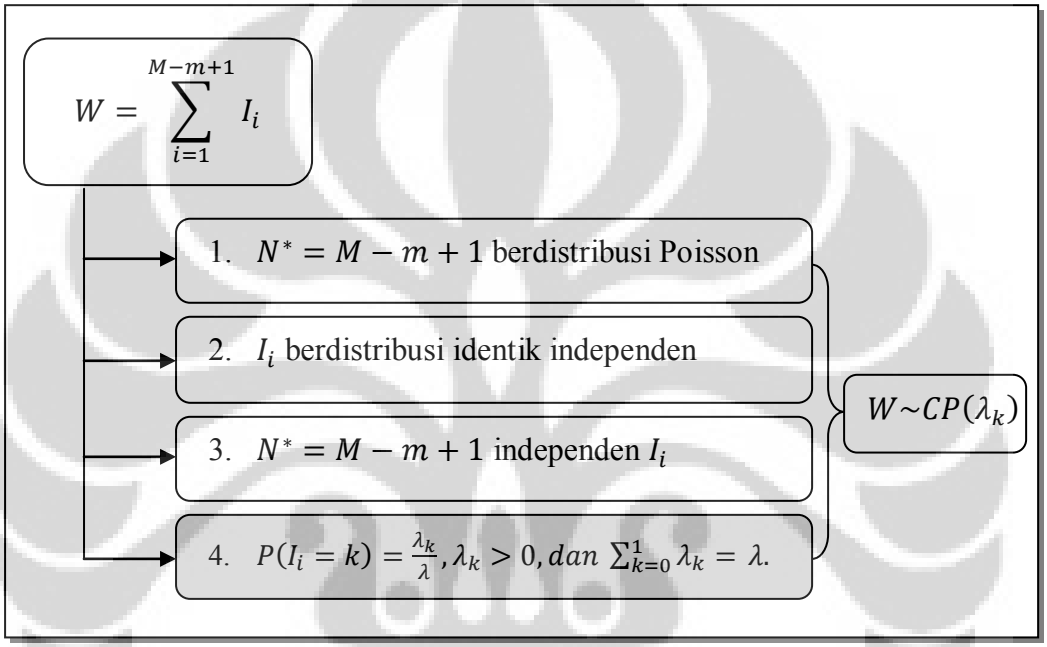
3.2.2 Sifat distribusi W

Pada pembahasan berikut akan dijelaskan bahwa W pada persamaan 3.5 berdistribusi *Compound Poisson*. W berdistribusi *Compound Poisson* apabila W memenuhi sifat:

1. $N^* = M - m + 1$ merupakan variabel acak berdistribusi Poisson.
2. I_i berdistribusi identik independen untuk $i = 1, 2, \dots, M - m + 1$.
3. $N^* = M - m + 1$ independen terhadap I_i.

Ketiga sifat tersebut merupakan sifat umum yang harus dipenuhi W agar dikatakan berdistribusi *Compound Poisson*. Namun berdasarkan Nuel (2006), harus ada satu tambahan sifat khusus dalam distribusi *Compound Poisson*nya dengan variabel acak I_i merupakan varibel acak diskrit, yaitu

4. $P(I_i = k) = \frac{\lambda_k}{\lambda}, \lambda_k > 0,$ dan $\sum_{k=0}^{1} \lambda_k = \lambda, \forall k \epsilon \mathbb{N}^*$



Gambar 3. 3 *Flowchart* sifat - sifat *W*

W merupakan banyaknya '*hit*' pada suatu blok ke - j dengan panjang M; *template B* dengan panjang m dalam barisan bilangan biner independen Be(p) dengan variabel acak I_i untuk $i = 1, 2, ... + I_{M-m+1}$, dimana

$$W = \sum_{i=1}^{M-m+1} I_i \text{ , dan}$$

$$I_i = \prod_{j=i}^{i+m-1} \varepsilon_j \text{ , } \quad 1 \leq i \leq M - m + 1, \text{dimana}$$

$$I_i = \begin{cases} 1 \text{ , jika terjadi '} hit' \\ 0 \text{ , jika tidak terjadi '} hit'. \end{cases}$$

1. $N^* = M - m + 1$ merupakan variabel acak berdistribusi Poisson.

   $N^* = M - m + 1$ merupakan variabel acak karena nilai dari N^* bergantung pada M yang dikeluarkan oleh perangkat lunak secara acak. Hal ini dikarenakan perangkat lunak yang mengeluarkan M dengan menggunakan suatu algoritma tertentu dengan distribusi Poisson di dalamnya. Barisan bilangan biner dibagi menjadi sebanyak N blok, dengan besar M tidak selalu sama tiap kali barisan tersebut dihasilkan.

   Kemudian berdasarkan definisinya $N^* = M - m + 1$ merupakan banyaknya pemeriksaan dengan *template* ukuran m dalam suatu blok dengan panjang M. Dikarenakan banyaknya pemeriksaan dalam blok yang *non-overlapping* adalah independen (karena pada kasus ini masing-masing blok dalam barisan bilangan biner tidak saling menutupi, serta banyaknya pemeriksaan dalam suatu blok tertentu tidak mempengaruhi blok lainnya), sehingga dapat dikatakan bahwa variabel acak $N^* = M - m + 1$ merupakan variabel acak yang berdistribusi Poisson. Pada kasus ini, banyaknya kejadian adalah banyaknya pemeriksaan dengan menggunakan *template* ukuran m dan selang intervalnya adalah blok sepanjang M tersebut.

2. I_i merupakan variabel acak yang memenuhi asumsi identik independen untuk $i = 1, 2, \dots, M - m + 1$.

   a. Identik

$$E(I_i) = E\left(\prod_{j=i}^{i+m-1} \varepsilon_j\right)$$

$$= E(\varepsilon_i . \varepsilon_{i+1} . \dots \varepsilon_{i+m-1})$$

$$= E(\varepsilon_i) . E(\varepsilon_{i+1}) . \dots E(\varepsilon_{i+m-1}), \text{ karena } \varepsilon_j \text{ dan } \varepsilon_k \text{ saling bebas untuk } j \neq k$$

$$= p . p . \dots p$$

$$= p^m \text{ , untuk } 1 \leq i \leq M - m + 1$$

Maka $I_i \sim \text{Be}(p^m)$ untuk setiap $i = 1,2,\dots, M-m+1$ sedemikian sehingga dapat dikatakan I_i identik.

$\therefore I_i$ identik

b. Independen

Akan dibuktikan variabel acak I_j dan I_k independen untuk $j \neq k$.

Secara umum:

Akan dibuktikan bahwa

$$\Pr(I_k = \text{'}hit\text{'} \mid I_j = \text{'}hit\text{'}) = \Pr(I_k = \text{'}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

dan

$$\Pr(I_k = \text{'tidak }hit\text{'} \mid I_j = \text{tidak '}hit\text{'}) = \Pr(I_k = \text{'tidak }hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

Pertama - tama perhatikan bentuk berikut:

$$\begin{aligned}
&\Pr(I_k = \text{'}hit\text{'} | I_j = \text{'}hit\text{'}) \\
&= \Pr(\varepsilon_k \dots \varepsilon_{j+m-1} \dots \varepsilon_{k+m-1} = 1 | \varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} \\
&= 1) \\
&= \Pr(\varepsilon_{j+m} \dots \varepsilon_{k+m-1} = 1) \\
&= \Pr(\varepsilon_{j+m} = 1) \Pr(\varepsilon_{j+m+1} = 1) \dots \Pr(\varepsilon_{k+m-1} \\
&= 1) \\
&= \left(\frac{1}{2}\right)^{k-j}
\end{aligned}$$

Kemudian,

$$\begin{aligned}
&\Pr(I_k = \text{'}hit\text{'} | I_j = \text{tidak '}hit\text{'}) \\
&= \Pr(\varepsilon_k \dots \varepsilon_{j+m-1} \dots \varepsilon_{k+m-1} = 1 | \varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} \\
&= 0)
\end{aligned}$$

$$= \Pr(\varepsilon_{j+m} \dots \varepsilon_{k+m-1} = 1)$$

$$= \Pr(\varepsilon_{j+m} = 1)\Pr(\varepsilon_{j+m+1} = 1) \dots \Pr(\varepsilon_{k+m-1} = 1)$$

$$= \left(\frac{1}{2}\right)^{k-j}$$

Sehingga terbukti bahwa

$$\Pr(I_k = \text{'}hit\text{'} \mid I_j = \text{'}hit\text{'}) = \Pr(I_k = \text{'}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

Kemudian perhatikan bentuk berikut,

$$\Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{'}hit\text{'}) = \Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

$$\Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{'}hit\text{'})$$

$$= \Pr(\varepsilon_k \dots \varepsilon_{j+m-1} \dots \varepsilon_{k+m-1} = 0 | \varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} = 1)$$

$$= \frac{\Pr(\varepsilon_k \dots \varepsilon_{j+m-1} \dots \varepsilon_{k+m-1} = 0, \varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} = 1)}{\Pr(\varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} = 1)}$$

$$= \frac{\Pr(\varepsilon_k \dots \varepsilon_{j+m-1} \dots \varepsilon_{k+m-1} = 0, \varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} = 1)}{\Pr(\varepsilon_j \dots \varepsilon_k \dots \varepsilon_{j+m-1} = 1)}$$

$$= \frac{\Pr(\varepsilon_{j+m} \dots \varepsilon_{k+m-1} = 0)}{\left(\frac{1}{2}\right)^m}$$

$$= \frac{1 - \Pr(\varepsilon_{j+m} \dots \varepsilon_{k+m-1} = 1)}{\left(\frac{1}{2}\right)^m}$$

$$= \frac{1 - \left(\left(\frac{1}{2}\right)^{k-j}\right)}{\left(\frac{1}{2}\right)^m}$$

$$\Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

$$= \frac{\Pr(I_k = \text{tidak '}hit\text{'}, I_j = \text{tidak '}hit\text{'})}{\Pr(I_j = \text{tidak '}hit\text{'})}$$

$$= \frac{1 - \Pr(I_k = \text{'}hit\text{' atau } I_j = \text{'}hit\text{'})}{1 - \Pr(I_j = \text{'}hit\text{'})}$$

$$= \frac{1 - [\Pr(I_k = \text{'}hit\text{'}) + \Pr(I_j = \text{'}hit\text{'}) - \Pr(I_k = \text{'}hit\text{'}, I_j = \text{'}hit\text{'})]}{1 - \Pr(I_j = \text{'}hit\text{'})}$$

$$= \frac{1 - \left[\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^m - \left(\frac{1}{2}\right)^{k-j}\right]}{1 - \left(\frac{1}{2}\right)^m}$$

$$= \frac{1 - \left(\left(\frac{1}{2}\right)^{k-j}\right)}{\left(\frac{1}{2}\right)^m}$$

Sehingga terbukti bahwa

$$\Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{'}hit\text{'}) = \Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

Karena telah dibuktikan bahwa

$$\Pr(I_k = \text{'}hit\text{'} \mid I_j = \text{'}hit\text{'}) = \Pr(I_k = \text{'}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

dan

$$\Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{'}hit\text{'}) = \Pr(I_k = \text{tidak '}hit\text{'} \mid I_j = \text{tidak '}hit\text{'})$$

$\therefore$ I_j dan I_k independen untuk $j \neq k$.

$\therefore$ variabel acak I_i berdistribusi identik independen.

3. Variabel acak $N^* = M - m + 1$ independen terhadap I_i.

   Jelas bahwa variabel acak $N^* = M - m + 1$ independen terhadap I_i. Berdasarkan definisi keduanya yaitu $N^* = M - m + 1$ merupakan banyaknya pemeriksaan yang dilakukan dalam suatu blok sedangkan I_i menyatakan '*hit*' atau tidaknya bagian dari blok tersebut yang sedang

diperiksa, terlihat jelas bahwa keduanya tidak saling mempengaruhi, sehingga dikatakan $N^* = M - m + 1$ independen terhadap I_i.

4. $P(I_i = k) = \frac{\lambda_k}{\lambda}$, $\lambda_k > 0$, dan $\sum_{k=0}^{1} \lambda_k = \lambda, \forall k \epsilon \mathbb{N}^*$

Karena telah dibuktikan bahwa I_i berdistribusi Bernoulli, maka dapat dituliskan:

$$P(I_i = k) = (p^m)^k (1 - p^m)^k, \quad k = 0, 1$$

Pilih $\lambda_k = \lambda(p^m)^k(1 - p^m)^k > 0$, dan $\sum_{k=0}^{1} \lambda_k = \lambda_0 + \lambda_1 = \lambda(p^m)^1 + \lambda(1 - p^m)^1 = \lambda$. Sehingga:

$$P(I_i = k) = (p^m)^k (1 - p^m)^k = \frac{\lambda_k}{\lambda}, \lambda_k > 0 \text{ dan } \sum_{k=0}^{1} \lambda_k = \lambda \quad k = 0, 1$$

Sehingga terbukti bahwa W berdistribusi *Compound Poisson*. Untuk selanjutnya akan dituliskan $W \sim CP(\lambda_k)$. Pada subbab selanjutnya akan dilakukan penurunan distribusi dari W dengan cara menurunkan p.d.f dari W

3.3 Penurunan Distribusi W

Berdasarkan subbab sebelumnya, telah dibuktikan bahwa W berdistribusi *Compound Poisson*, dinotasikan $W \sim CP(\lambda_k)$, dengan parameter λ_k, dengan $\lambda_k > 0$ dan $\sum_{k=0}^{1} \lambda_k = \lambda$, λ_k adalah parameter untuk setiap I_i, yang menyatakan 'bobot' bahwa $I_1, I_2, \ldots, I_m$ masuk ke kelas, dimana:

$$W = \sum_{i=1}^{N^*} I_i = I_1 + I_2 + \ldots + I_{N^*}, \text{ dimana } N^* = M - m + 1$$

dengan N^* berdistribusi Poisson dan independen terhadap I_i, dimana I_i berdistribusi identik dan independen, dan

$$P(I_i = k) = \frac{\lambda_k}{\lambda}, \lambda_k > 0, \text{ dan } \sum_{k=0}^{1} \lambda_k = \lambda, \forall k \epsilon \mathbb{N}^*.$$

Selanjutnya akan dilihat bentuk distribusi dari W. Untuk memudahkan pembahasan maka pada penjelasan dibawah ini akan digunakan ilustrasi berdasarkan Kaas, Govaerts, Dhaene, & Denuit (2002), yaitu masalah total klaim dari sejumlah orang (dalam kasus ini, total klaim yang dibahas untuk total klaim yang diskrit), dimana jumlah orang yang akan mengajukan klaim mengikuti distribusi Poisson dan besarnya klaim dari masing-masing orang memiliki distribusi tertentu.

Karena $W \sim CP(\lambda_k)$, dengan $\sum_{k=0}^{1} \lambda_k = \lambda$, maka berdasarkan Lemma 2.1, $\forall k \epsilon \mathbb{N}^*$,

$$P(W = i) = \sum_{l=1}^{i} e^{-\lambda} \frac{\lambda^l}{l!} \sum_{k_1,\dots,k_l \in \mathbb{N}^*} \mathrm{I}_{\{k_1+\cdots+k_l=i\}} \frac{\lambda_{k_1} \times \dots \times \lambda_{k_l}}{\lambda^l}$$

dan $P(W = 0) = e^{-\lambda}$.

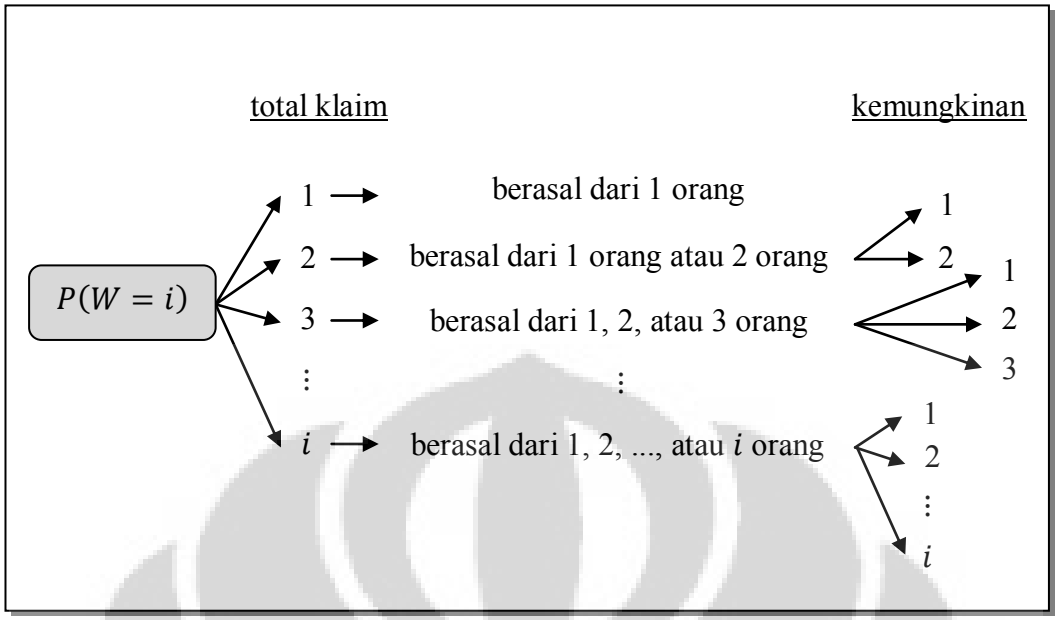
Dari bentuk tersebut dapat digambarkan sebagai berikut:

$e^{-\lambda} \frac{\lambda^l}{l!}$: menyatakan peluang l orang yang mengajukan klaim.

$\mathrm{I}_{\{k_1+k_2\dots+k_l=i\}}$: indikator bahwa nilai klaim sebesar i berasal dari l orang.

$\frac{\lambda_{k_1} \times \dots \times \lambda_{k_l}}{\lambda^l}$: kemungkinan bahwa klaim sebesar i berasal dari l orang.

Sehingga secara keseluruhan, $P(W = i)$ merupakan probabilitas total klaim yang diajukan sebesar i. Dalam hal ini, orang yang mengajukan klaim tersebut bisa berasal dari 1 orang, atau 2 orang, hingga i orang (dilihat berdasarkan nilai klaim dari kemungkinan orang yang mengajukan). Untuk lebih mudahnya akan diberikan gambaran sebagai berikut:



Gambar 3. 4 *Chart* kemungkinan total klaim W

Untuk $i = 1$,

$$P(W = 1) = e^{-\lambda}\frac{\lambda}{1!}\left(\mathrm{I}_{\{k_1=1\}}\frac{\lambda_{k_1}}{\lambda}\right)$$

Artinya probabilitas total klaim yang berasal dari 1 orang mengajukan adalah total nilai klaim yang diperoleh dari 1 orang tersebut saja.

Untuk $i = 2$,

$$P(W = 2) = e^{-\lambda}\frac{\lambda^2}{2!}\left(\mathrm{I}_{\{k_1=2\}}\frac{\lambda_{k_1}}{\lambda} + \mathrm{I}_{\{k_1+k_2=2\}}\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)$$

Artinya probabilitas total klaim yang berasal dari 2 orang yang mengajukan adalah total nilai klaim bisa berasal dari 1 orang saja atau total nilai klaim bisa berasal dari kedua orang tersebut. Untuk nilai $i = 3, 4, ...,$ dan selanjutnya, probabilitasnya akan seperti yang digambarkan di atas.

Kemudian akan dibuktikan bahwa W memiliki bentuk distribusi khusus dari distribusi *Compound Poisson*. Perhatikan bentuk dari W,

$$W = \sum_{i=1}^{M-m+1} I_i,$$

pada pembahasan di awal subbab 3.2.2, telah diketahui bahwa bentuk dari W merupakan penjumlahan dari variabel acak I_i yang berdistribusi *Compound Poisson*.

Pertama - tama perhatikan barisan bilangan biner dalam satu blok:

$$\varepsilon_i \sim \text{Be}(p) \text{ untuk } 0 \leq p \leq 1$$

Karena selanjutnya akan diturunkan distribusi dari W dibawah asumsi bahwa barisan bilangan biner yang diberikan adalah acak, maka $p = \frac{1}{2}$. Karena peluang keluarnya bilangan 1 dan 0 sama.

Misal untuk *template B* dengan $m = 2$, akan terdapat $2^2 = 4$ kemungkinan, yaitu $00, 01, 10, 11$. Maka probabilitas akan didapatkan kejadian misal untuk $11 = \frac{1}{2^2} = \frac{1}{4}$. Sehingga untuk *template B* dengan panjang m, probabilitas akan terjadi '*hit*' adalah $\frac{1}{2^m}$, dengan kata lain $p = \frac{1}{2^m}$.

- Pada suatu *template*, hanya terjadi 2 kemungkinan, '*hit*' atau tidak '*hit*'. Sebut '*hit*' sebagai sukses, dengan probabilitas akan terjadi sukses adalah $\frac{1}{2^m}$.
- Pada suatu blok dengan panjang M, ada sebanyak $M - m + 1$ pemeriksaan dengan menggunakan *template* ukuran m. Telah diketahui pula bahwa $N^* = M - m + 1$ merupakan suatu variabel acak. $N^* = M - m + 1$ disebut banyaknya percobaan atau pengecekan.

Sehingga banyaknya '*hit*' dalam satu blok tertentu berdistribusi binomial, dengan parameter $p = \frac{1}{2^m}$ adalah probabilitas terjadinya '*hit*' dan $M - m + 1$ adalah banyaknya pengecekan. Lebih lanjut distribusi dari banyaknya '*hit*' ini dapat didekati dengan distribusi Poisson, dengan parameter $\lambda = (M - m + 1)\frac{1}{2^m}$.

Pendekatan dari distribusi binomial ke distribusi Poisson tersebut dapat dilakukan apabila banyaknya percobaan atau pengecekan yaitu $M - m + 1$ bernilai besar, dan nilai p sangat kecil. Berdasarkan Rukhin, et al (2001), nilai m

yang dipilih sebaiknya adalah $m = 9$ atau $m = 10$, maka nilai $p = \frac{1}{2^m}$ akan bernilai sangat kecil. Sehingga didapat parameter $\lambda = \frac{M-m+1}{2^m} > 0$.

Untuk melihat bentuk khusus dari distribusi W, pada pembahasan setelah ini parameter $\lambda > 0$ dikatakan sebagai parameter untuk bagian Poisson.

Kemudian pada akhir subbab 3.2.2 telah diperlihatkan bahwa W yang berdistribusi *Compound Poisson* memiliki parameter λ_k dimana

$$\lambda_k = \lambda(\theta)^k(1-\theta)^k, \qquad k = 0,1.$$

dimana θ merupakan parameter dari distribusi Bernoulli, dengan $0 \leq \theta = \frac{1}{2} \leq 1$. Sama seperti λ sebagai parameter untuk bagian Poisson, parameter θ untuk selanjutnya akan dikatakan sebagai parameter untuk bagian Geometrik.

Berdasarkan definisi pada subbab 2.2.5, jika $W \sim CP(\lambda_k)$, dengan $\lambda_k = \lambda(\theta)^k(1-\theta)^k,\ k = 0,1$, λ sebagai parameter untuk bagian Poisson, dan θ sebagai parameter untuk bagian Geometrik, maka dikatakan W memiliki suatu distribusi khusus dari *Compound Poisson* yaitu W berdistribusi *Geometric Poisson* (*Pólya-Aeppli*). Dinotasikan dengan $W \sim GP(\lambda, \theta)$.

Kemudian berdasarkan Proposisi 2.1, karena $W \sim GP(\lambda, \theta)$, maka untuk setiap $i \epsilon \mathbb{N}^*$

$$P(W = i) = \sum_{l=1}^{i} e^{-\lambda} \frac{\lambda^l}{l!} (1-\theta)^{i-l} (\theta)^l \binom{i-1}{l-1}$$

dimana $\binom{i}{l}$ merupakan koefisien binomial, dan $P(W = 0) = e^{-\lambda}$.

Terdapat bentuk lain untuk probabilitas W seperti yang dipaparkan oleh Johnson, Kotz, & Kemp, 1996, dimana apabila W berdistribusi *Pólya-Aeppli*, dengan parameter λ sebagai parameter untuk bagian Poisson, dan parameter p sebagai parameter untuk bagian Geometriknya, maka dapat dituliskan:

$$W \sim \text{Poisson}(\lambda) \bigvee \text{Geometrik}(p)$$

dimana nilai parameter λ didefinisikan sebagai $\lambda = \frac{\eta}{p}$. Telah ketahui bahwa $\lambda = \frac{M-m+1}{2^m} > 0$ sebagai parameter untuk bagian Poisson, sedangkan untuk parameter geometriknya, $p = \frac{1}{2}$, karena kemungkinan keluarnya bilangan biner 1 atau 0 saling bebas (masing - masing bilangan biner memiliki peluang keluar sama). Dapat direpresentasikan sebagai berikut:

$$W \sim \text{Poisson}\left(\frac{\eta}{p}\right) \bigvee \text{Geometrik}(p).$$

Kemudian akan didefinisikan suatu $q = 1 - p$, dengan $p = \frac{1}{2}$. Dengan $\lambda = \frac{\eta}{p}$ maka nilai $\eta = p\lambda = \frac{\lambda}{2}$. Sehingga berdasarkan Johnson, Kotz, & Kemp, 1996, bentuk distribusi dari W yaitu:

$$P(W = 0) = e^{-\eta}$$

$$P(W = i) = e^{-\eta} p^i \sum_{l=1}^{i} \binom{i-1}{l-1} \frac{\left(\frac{\eta q}{p}\right)^l}{l!}$$

dengan

$p = \frac{1}{2}$, maka $q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$.

Sehingga untuk $\eta = \frac{\lambda}{2}$ bentuk distribusinya dapat ditulis:

$$P(W = i) = e^{-\eta} \left(\frac{1}{2}\right)^i \sum_{l=1}^{i} \binom{i-1}{l-1} \frac{\left(\frac{\eta\, 1/2}{1/2}\right)^l}{l!}$$

$$= \frac{e^{-\eta}}{2^i} \sum_{l=1}^{i} \binom{i-1}{l-1} \frac{\eta^l}{l!}$$

Sekarang akan dilakukan transformasi pada bentuk distribusi ini dengan menggunakan *Confluent Hypergeometric Function* (*Kummer*'s). Hal ini dilakukan agar bentuk distribusinya menjadi lebih singkat dan juga dalam kenyataannya pada mesin akan jauh lebih mudah jika sudah terdapat bentuk fungsi yang sudah

lebih familiar, yaitu *Confluent Hypergeometric Function* (*Kummer*'s). Berdasarkan Proposisi 2.2, diketahui bahwa:

$$\sum_{l=1}^{i}\binom{i-1}{l-1}\frac{\eta^{l-1}}{l!}=\Phi(-i+1,2,-\eta)$$

dimana $\Phi(-i+1,2,-\eta)=e^{-\eta}\Phi(i+1,2,\eta)$.

Sehingga bentuk distribusi dari W dengan menggunakan *Confluent Hypergeometric Function* (*Kummer's M*), didapat yaitu:

$$P(W=i)=\frac{e^{-\eta}}{2^i}\sum_{l=1}^{i}\binom{i-1}{l-1}\frac{\eta^{l}}{l!}$$

$$=\frac{\eta e^{-\eta}}{2^i}\sum_{l=1}^{i}\binom{i-1}{l-1}\frac{\eta^{l-1}}{l!}$$

$$=\frac{\eta e^{-\eta}}{2^i}\Phi(-i+1,2,-\eta)$$

$$=\frac{\eta e^{-\eta}}{2^i}e^{-\eta}\Phi(i+1,2,\eta)$$

$$=\frac{\eta e^{-2\eta}}{2^i}\Phi(i+1,2,\eta)$$

Didapat bentuk distribusi dari W dengan menggunakan *Confluent Hypergeometric Function* (*Kummer's*), dimana $i\geq 1$ dengan $\eta=\lambda/2$

$$\pi_i=P(W=i)=\frac{\eta e^{-2\eta}}{2^i}\Phi(i+1,2,\eta)$$

dimana:

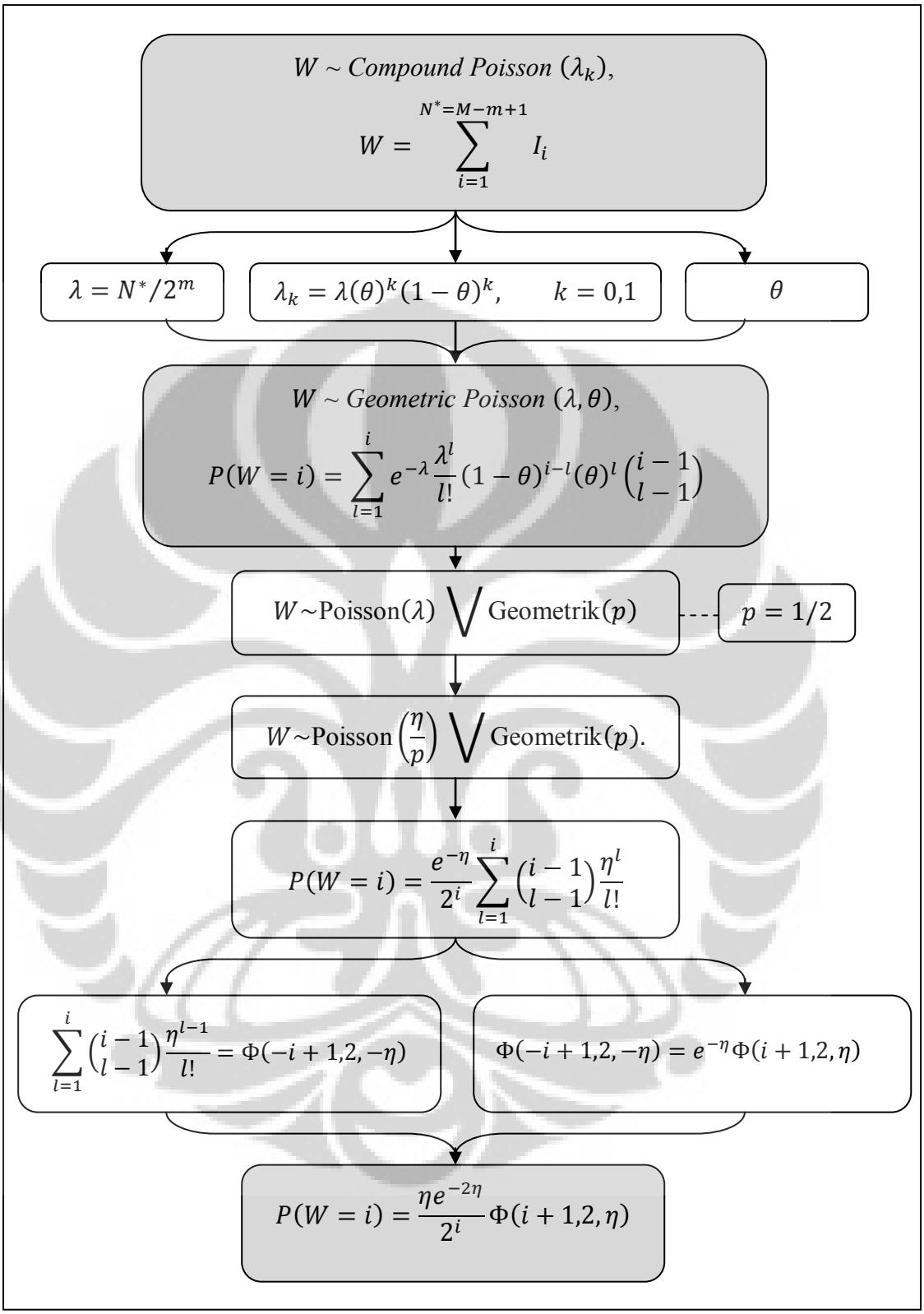
$\lambda=\dfrac{M-m+1}{2^m}>0$, dimana M merupakan panjang blok, dan

m merupakan panjang *template*.

$\Phi(\mathrm{a},\mathrm{b},\mathrm{z})=$ *Confluent Hypergeometric Function* (*Kummer's*).

$\pi_i = P(W = i)$, menyatakan probabilitas suatu blok (diantara blok - blok dalam barisan) yang memuat i '*hit*'.

Dari pembahasan ini, proses penurunan distribusi W dapat diringkas dalam *flowchart* berikut:



Gambar 3. 5 *Flowchart* Penurunan Distribusi *W*

BAB 4
ILUSTRASI

Berikut ini merupakan contoh dari penentuan kerandoman barisan bilangan biner pada metode Overlapping *Template* Matching Test.

Ilustrasi:

Untuk barisan bilangan biner

$\varepsilon =$ 1011101111**00101101**00111001011**01111000**0101101001 → $n = 50$

Blok 1 **Blok 2** **Blok 3** **Blok 4** **Blok 5** → $N = 5$

dengan $M = 10$ (panjang tiap blok) dan $N = 5$ (banyak blok). Kemudian pilih $m = 2$ dan $B = 11$, didapat:

Blok 1

Tabel 4. 1 Tabel evaluasi blok 1

Posisi *bit*	*Bit*	Akumulasi untuk $B = 11$
$1-2$	10	0
$2-3$	01	0
$3-4$	11	'*hit*' ke -1
$4-5$	11	'*hit*' ke -2
$5-6$	10	0
$6-7$	01	0
$7-8$	11	'*hit*' ke -3
$8-9$	11	'*hit*' ke -4
$9-10$	11	'*hit*' ke -5

Terjadi 5 *hit* dengan $B = 11$, maka $\boldsymbol{v_5 = 1}$, $v_0 = 0, v_1 = 0, v_2 = 0, v_3 = 0$, dan $v_4 = 0$.

Blok 2

Tabel 4. 2 Tabel evaluasi blok 2

Posisi *bit*	*Bit*	Banyak kejadian dengan $B = 11$
$11 - 12$	00	0
$12 - 13$	01	0
$13 - 14$	10	0
$14 - 15$	01	0
$15 - 16$	11	'*hit*' ke -1
$16 - 17$	10	0
$17 - 18$	01	0
$18 - 19$	10	0
$19 - 20$	00	0

Terjadi 1 *hit* dengan $B = 11$, maka $\boldsymbol{v_1 = 1}, v_0 = 0, v_2 = 0, v_3 = 0, v_4 = 0$, dan $v_5 = 1$.

Blok 3

Tabel 4. 3 Tabel evaluasi blok 3

Posisi *bit*	*Bit*	Banyak kejadian dengan $B = 11$
$21 - 22$	01	0
$22 - 23$	11	'*hit*' ke -1
$23 - 24$	11	'*hit*' ke -2
$24 - 25$	10	0
$25 - 26$	00	0
$26 - 27$	01	0
$27 - 28$	10	0
$28 - 29$	01	0

$29-30$	11	'*hit*' ke -3

Terjadi 3 *hit* dengan $B = 11$, maka $\boldsymbol{v_3 = 1}, v_0 = 0, v_1 = 1, v_2 = 0, v_4 = 0$, dan $v_5 = 1$.

Blok 4

Tabel 4. 4 Tabel evaluasi blok 4

Posisi *bit*	*Bit*	Banyak kejadian dengan $B = 11$
$31-32$	10	0
$32-33$	01	0
$33-34$	11	'*hit*' ke -1
$34-35$	11	'*hit*' ke -2
$35-36$	11	'*hit*' ke -3
$36-37$	11	'*hit*' ke -4
$37-38$	10	0
$38-39$	00	0
$39-40$	00	0

Terjadi 4 *hit* dengan $B = 11$, maka $\boldsymbol{v_4 = 1}, v_0 = 0, v_1 = 1, v_2 = 0, v_3 = 1$, dan $v_5 = 1$.

Blok 5

Tabel 4. 5 Tabel evaluasi blok 5

Posisi *bit*	*Bit*	Banyak kejadian dengan $B = 11$
$41-42$	01	0
$42-43$	10	0
$43-44$	01	0

$44-45$	11	'*hit*' ke -1
$45-46$	10	0
$46-47$	01	0
$47-48$	10	0
$48-49$	00	0
$49-50$	01	0

Terjadi 1 *hit* dengan $B = 11$, maka $\boldsymbol{v_1 = 2}, v_0 = 0, v_2 = 0, v_3 = 1, v_4 = 1$, dan $v_5 = 1$.

Sehingga setelah pengecekan pada tiap blok, didapat:

$$v_0 = 0, v_1 = 2, v_2 = 0, v_3 = 1, v_4 = 1, \text{ dan } v_5 = 1$$

Kemudian hitung nilai λ dan η untuk mendapatkan nilai π_i:

$$\lambda = \frac{M - m + 1}{2^m} = \frac{10 - 2 + 1}{2^2} = 2.25$$

$$\eta = \frac{\lambda}{2} = \frac{2.25}{2} = 1.125$$

untuk $i \geq 1$ dengan $\eta = {}^{\lambda}/_{2}$,

$$\pi_i = P(W = i) = \frac{e^{-\eta}}{2^i} \sum_{l=1}^{i} \binom{i-1}{l-1} \frac{n^l}{l!} = \frac{\eta e^{-2\eta}}{2^u} \Phi(i+1, 2, \eta)$$

$$\pi_0 = P(W = 0) = e^{-\eta}$$

dimana

$l = 1, 2, \ldots i$

$\Phi(a, b, z) = 1 + \frac{az}{b} + \frac{(a)_2 z^2}{(b)_2 2!} + \ \ldots + \frac{(a)_n z^n}{(b)_n n!} + \ \ldots$, dimana

$(a)_n = a(a+1)(a+2)\ldots(a+n-1)$

$(a)_0 = 1.$

maka

$$\pi_0 = P(W = 0)$$

$$= e^{-\eta}$$

$$= e^{-1.125}$$

$$= 0.324652$$

$$\pi_1 = P(W = 1)$$

$$= \frac{e^{-\eta}}{2}\eta$$

$$= \frac{e^{-1.125}}{2}1.125$$

$$= 0.182617$$

$$\pi_2 = P(W = 2)$$

$$= \frac{e^{-\eta}}{2^2}\sum_{l=1}^{2}\binom{2-1}{l-1}\frac{\eta^l}{l!}$$

$$= \frac{e^{-\eta}}{2^2}\left(\frac{\eta}{1!}+\frac{\eta^2}{2!}\right)$$

$$= \frac{e^{-\eta}}{2^2}\eta\left(1+\frac{\eta}{2}\right)$$

$$= \frac{\eta e^{-\eta}}{8}(2+\eta)$$

$$= \frac{1.125e^{-1.125}}{8}(3.125)$$

$$= 0.14267$$

$$\pi_3 = P(W = 3)$$

$$= \frac{e^{-\eta}}{2^3}\sum_{l=1}^{3}\binom{3-1}{l-1}\frac{\eta^l}{l!}$$

$$= \frac{e^{-\eta}}{8}\left[\binom{2}{0}\frac{\eta}{1!} + \binom{2}{1}\frac{\eta^2}{2!} + \binom{2}{2}\frac{\eta^3}{3!}\right]$$

$$= \frac{\eta e^{-\eta}}{8}\left(1 + \eta + \frac{\eta^2}{6}\right)$$

$$= \frac{1.125e^{-1.125}}{8}\left(1 + 1.125 + \frac{1.125^2}{6}\right)$$

$$= 0.106645$$

$$\pi_4 = P(W = 4)$$

$$= \frac{e^{-\eta}}{2^4}\sum_{l=1}^{4}\binom{4-1}{l-1}\frac{\eta^l}{l!}$$

$$= \frac{e^{-\eta}}{16}\left[\binom{3}{0}\frac{\eta}{1!} + \binom{3}{1}\frac{\eta^2}{2!} + \binom{3}{2}\frac{\eta^3}{3!} + \binom{3}{3}\frac{\eta^4}{4!}\right]$$

$$= \frac{\eta e^{-\eta}}{16}\left(1 + \frac{3\eta}{2} + \frac{\eta^2}{2} + \frac{\eta^3}{24}\right)$$

$$= \frac{1.125e^{-1.125}}{16}\left(1 + \frac{3(1.125)}{2} + \frac{1.125^2}{2} + \frac{1.125^3}{24}\right)$$

$$= 0.077147$$

$$\pi_5 = P(W = 5)$$

$$= \frac{e^{-\eta}}{2^5}\sum_{l=1}^{5}\binom{5-1}{l-1}\frac{\eta^l}{l!}$$

$$= \frac{e^{-\eta}}{32}\left[\binom{4}{0}\frac{\eta}{1!} + \binom{4}{1}\frac{\eta^2}{2!} + \binom{4}{2}\frac{\eta^3}{3!} + \binom{4}{3}\frac{\eta^4}{4!} + \binom{4}{4}\frac{\eta^5}{5!}\right]$$

$$= \frac{\eta e^{-\eta}}{32}\left(1 + 2\eta + 2\eta^2 + \frac{\eta^3}{6} + \frac{\eta^4}{120}\right)$$

$$= 0.166269$$

Sehingga didapat:

$$\pi_0 = 0.324625, \pi_1 = 0.182617, \pi_2 = 0.14267, \pi_3 = 0.106645, \pi_4 = 0.077147 \;\&\; \pi_5 = 0.166269.$$

Dengan hipotesis :

H_0 : Barisan biner random

$P(W = 0) = \pi_0 = e^{-\eta}$

$$P(W = i) = \pi_i = \frac{e^{-\eta}}{2^i}\sum_{l=1}^{i}\binom{i-1}{l-1}\frac{\eta^l}{l!} = \frac{\eta e^{-2\eta}}{2^i}\Phi(i+1, 2, \eta); \text{untuk } i \geq 1, \eta = \frac{\lambda}{2},$$

H_1 : Tidak demikian

Tingkat signifikasi : $\alpha = 0.01$

Statistik uji :

$$\chi^2 = \sum_{i=0}^{5}\frac{(v_i - N\pi_i)^2}{N\pi_i}$$

Aturan Keputusan : H_0 ditolak jika $\chi^2 \geq \chi^2_{(5)}$

Setelah nilai dari v_i dan π_i ditemukan, kemudian hitung nilai dari statistik uji *chi-square*:

$$\chi^2 = \sum_{i=0}^{5} \frac{(v_i - N\pi_i)^2}{N\pi_i}$$

$$= \frac{\left(0 - 5(0.324652)\right)^2}{5(0.324652)} + \frac{\left(2 - 5(0.182617)\right)^2}{5(0.182617)}$$

$$+ \frac{\left(0 - 5(0.14267)\right)^2}{5(0.14267)} + \frac{\left(1 - 5(0.106645)\right)^2}{5(0.106645)}$$

$$+ \frac{\left(1 - 5(0.077147)\right)^2}{5(0.077147)} + \frac{\left(1 - 5(0.166269)\right)^2}{5(0.166269)}$$

$$= 3.167729.$$

Keputusan : Karena $\chi^2 = 3.167729 < 15.1 = \chi^2_{(5)}$, berarti H_0 tidak ditolak.

Kesimpulan :

Dengan tingkat signifikasi $\alpha = 0.01$, karena $\chi^2 = 3.167729 < 15.1 = \chi^2_{(5)}$ maka H_0 tidak ditolak, sehingga dapat disimpulkan bahwa barisan biner tersebut ***random***.

BAB 5
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dalam pembahasan tugas akhir ini didapat kesimpulan bahwa penentuan distribusi dari banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test* didapat melalui beberapa tahapan yaitu, pertama suatu barisan bilangan biner ε dengan panjang n, dibagi menjadi N blok dengan panjang masing - masing blok adalah M dengan syarat $n \geq MN$. Kemudian gunakan *template* B dengan panjang m untuk menguji tiap - tiap blok.

Setelah itu definisikan I_i sebagai variabel acak yang menyatakan '*hit*' atau tidaknya bagian dari barisan suatu blok dengan menggunakan *template* B. Kemudian definisikan variabel acak W untuk suatu blok yang merepresentasikan banyaknya '*hit*' pada blok tersebut apabila menggunakan B sebagai *template* menguji tiap blok tersebut.

Kemudian dibuktikan variabel acak W berdistribusi *Compound Poisson*, lebih khusus lagi dibuktikan bahwa variabel acak W berdistribusi *Geometric Poisson* berdasar sifat yang dimiliki variabe acak W. Setelah didapat bentuk distribusi dari W yaitu *Geometric Poisson*, lakukan transformasi dengan menggunakan *Confluent Hypergeometric Function* (*Kummer's*). Sehingga pada akhirnya didapat bentuk distribusi dari variabel acak W yang merepresentasikan banyaknya '*hit*' kerandoman barisan bilangan biner pada metode *Overlapping Template Matching Test*.

5.2 Saran

Saran yang perlu diperhatikan adalah

1. Penurunan distribusi kerandoman barisan bilangan biner yang dilakukan pada tugas akhir ini menggunakan *template* dengan panjang m dan B untuk barisan dari '1', yaitu $B = 11$, $B = 1111$, dst.). Untuk pembahasan lebih lanjut perlu juga diperhatikan apabila akan digunakan *template* dengan panjang m dan B untuk barisan '0' dan '1' misal $B = 01$, $B = 001$.
2. Tugas akhir ini bertujuan untuk menurunkan distribusi kerandoman barisan bilangan biner dengan metode *Overlapping Template Matching Test*. Untuk lebih lanjut dapat dilakukan penurunan distribusi kerandoman barisan bilangan biner, dengan metode uji lain seperti metode *Non-overlapping Template Matching Test*, *Lempel-Ziv Test*, *Linear Complexity Test*, dan *Approximate Entropy Test*.
3. Untuk lebih lanjut dapat dilakukan perbandingan kekurangan maupun kelebihan antara metode *Overlapping Template Matching Test* dengan metode-metode lainnya dalam menguji kerandoman barisan bilangan biner.

DAFTAR PUSTAKA

Abramowitz, M., & Stegun, I. A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Washington, D.C.: Wiley-Interscience Publication.

Haahr, M. (1998-2011). *Randomness: RANDOM.ORG*. Dipetik Juli 4, 2011, dari RANDOM.ORG Web site: http://www.random.org/randomness/

Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics.* New Jersey: Prentice-Hall.

Johnson, N. L., Kotz, S., & Kemp, A. W. (1996). *Univariate Discrete Distributions.* New York: Wiley-Interscience Publication, 2nd ed.

Kaas, R., Govaerts, M., Dhaene, J., & Denuit, M. (2002). *Modern Actuarial Risk Theory.* Boston/ Dordrecht/ London: Kluwer Academic Publishers.

Nuel, G. (2006). *Cummulative Distribution Function of a Geometric Poisson Distribution.* Paris: Laboratoire Statistique & Génome Press.

Otniel. (2011). *Pembangkit Bilangan Acak dengan Memanfaatkan Fenomena Fisis.* Bandung: Program Studi Teknik Informatika - Institut Teknologi Bandung.

Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., et al. (2001). *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications.* Springfield: National Institute of Standards and Technology (NIST).

LAMPIRAN

Lampiran 1 Pembuktian Lemma 2.1

Diketahui

$$N = \sum_{m=1}^{M} K_m = K_1 + K_2 + \cdots + K_M$$

dimana $M \sim P(\lambda)$ independen terhadap K_m, K_m berdistribusi identik dan independen dengan

$$P(K = k) = \frac{\lambda_k}{\lambda} \qquad \forall k \epsilon \mathbb{N}^*$$

Untuk $\lambda_k > 0$, $\sum_{k=1}^{\infty} \lambda_k = \lambda$.

Untuk $n = 0$

$$P(N = 0) = P\left(\sum_{m}^{M} K_m = 0\right)$$

$$= \left(e^{-\lambda} \frac{\lambda^0}{0!}\right)$$

$$= e^{-\lambda}$$

$e^{-\lambda}$ merupakan probabilitas untuk menghasilkan $n = 0$. Probabilitas memilih K adalah 1, karena tidak ada K yang dapat dipilih.

Untuk $n = 1$

$$P(N = 1) = P\left(\sum_{m=1}^{M} K_m = 1\right)$$

$$= \sum_{m=1}^{M} P(K_m = 1)$$

$$= \left(e^{-\lambda}\frac{\lambda}{1!}\right)\left(\frac{\lambda_{k_1}}{\lambda}\right)$$

$\left(e^{-\lambda}\frac{\lambda}{1!}\right)$ merupakan probabilitas untuk membentuk $n = 1$ dari satu buah K (bagian Poissonnya), dan $\frac{\lambda_{k_1}}{\lambda}$ merupakan probabilitas memilih K mana yang digunakan untuk membentuk $n = 1$. Karena K identik maka $\frac{\lambda_{k_1}}{\lambda}$ dapat digunakan untuk setiap K yang mungkin. Untuk mempermudah, dapat ditulis dengan menggunakan sebuah indikator $I_{\{k_1+k_2+\cdots+k_m=1\}}$ sehingga

$$= \left(e^{-\lambda}\frac{\lambda}{1!}\right)\left(I_{\{k_1=1\}}\frac{\lambda_{k_1}}{\lambda}\right)$$

indikator $I_{\{k_1=1\}}$ menyatakan salah satu K yang dapat digunakan adalah bernilai 1.

Untuk $n = 2$

$$P(N = 2) = P\left(\sum_{m=1}^{M} K_m = 2\right)$$

$$= \sum_{m=1}^{M} P(K_m = 2) + \sum_{i\neq j}^{M}\sum_{j}^{M} P(K_i + K_j = 2)$$

$$= e^{-\lambda}\frac{\lambda}{1!}\left(I_{\{k_1=2\}}\frac{\lambda_{k_1}}{\lambda}\right) + e^{-\lambda}\frac{\lambda^2}{2!}\left[I_{\{k_1+k_2=2\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)\right]$$

$\left(e^{-\lambda}\frac{\lambda}{1!}\right)$ merupakan probabilitas untuk membentuk $n = 2$ dari satu buah K (bagian Poissonnya), dan $\left(I_{\{k_1=2\}}\frac{\lambda_{k_1}}{\lambda}\right)$ merupakan probabilitas memilih satu K mana yang dapat digunakan untuk membentuk $n = 2$, indikator $I_{\{k_1=2\}}$ menyatakan salah satu K yang dapat digunakan adalah bernilai 2.

Sedangkan untuk $\left(e^{-\lambda}\frac{\lambda^2}{2!}\right)$ merupakan probabilitas untuk membentuk $n = 2$ dari dua buah K (bagian Poissonnya), dan $\left(I_{\{k_1+k_2=2\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)\right)$ merupakan probabilitas memilih dua K mana yang digunakan untuk membentuk $n = 2$, indikator $I_{\{k_1+k_2=2\}}$ menyatakan dua K tersebut yang dapat digunakan menghasilkan jumlah nilai $n = 2$ (misal $K_1 = 1$ dan $K_2 = 1$).

Untuk $n = 3$

$$P(N=3) = P\left(\sum_{m=1}^{M} K_m = 3\right)$$

$$= \sum_{m=1}^{M} P(K_m = 3) + \sum_{i\neq j}^{M}\sum_{j}^{M} P\left(K_i + K_j = 3\right) + \sum_{i\neq j,k}^{M}\sum_{j\neq k}^{M}\sum_{k}^{M} P\left(K_i + K_j + K_k = 3\right)$$

$$= e^{-\lambda}\frac{\lambda}{1!}\left(I_{\{k_1=3\}}\frac{\lambda_{k_1}}{\lambda}\right) + e^{-\lambda}\frac{\lambda^2}{2!}\left[I_{\{k_1+k_2=3\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)\right] + e^{-\lambda}\frac{\lambda^3}{3!}\left[I_{\{k_1+k_2+k_3=3\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}\times\lambda_{k_3}}{\lambda^3}\right)\right]$$

$\left(e^{-\lambda}\frac{\lambda}{1!}\right)$ merupakan probabilitas untuk membentuk $n = 3$ dari satu buah K (bagian Poissonnya), dan $\left(I_{\{k_1=3\}}\frac{\lambda_{k_1}}{\lambda}\right)$ merupakan probabilitas memilih satu K mana yang dapat digunakan untuk membentuk $n = 3$, indikator $I_{\{k_1=3\}}$ menyatakan salah satu K yang dapat digunakan adalah bernilai 3.

Kemudian untuk $\left(e^{-\lambda}\frac{\lambda^2}{2!}\right)$ merupakan probabilitas untuk membentuk $n = 3$ dari dua buah K (bagian Poissonnya), dan $\left(I_{\{k_1+k_2=3\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)\right)$ merupakan probabilitas memilih dua K mana yang digunakan untuk membentuk $n = 3$, indikator $I_{\{k_1+k_2=3\}}$ menyatakan dua K tersebut yang dapat digunakan menghasilkan jumlah nilai $n = 3$ (misal $K_1 = 1$ dan $K_2 = 2$, atau $K_1 = 2$ dan $K_2 = 1$).

Selanjutnya untuk $\left(e^{-\lambda}\frac{\lambda^3}{3!}\right)$ merupakan probabilitas untuk membentuk $n = 3$ dari tiga buah K (bagian Poissonnya), dan $\left(I_{\{k_1+k_2+k_3=3\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}\times\lambda_{k_3}}{\lambda^3}\right)\right)$ merupakan probabilitas memilih tiga K mana yang digunakan untuk membentuk $n = 3$, indikator $I_{\{k_1+k_2+k_3=3\}}$ menyatakan tiga K tersebut yang dapat digunakan menghasilkan jumlah nilai $n = 3$ (misal $K_1 = 1, K_2 = 1$, dan $K_3 = 1$).

Sehingga secara umum dapat ditulis

$$P(N = n) = P\left(\sum_{m=1}^{M} K_m = n\right)$$

$$= \sum_{j_1=1}^{M} P(K_{j_1} = n) + \sum_{j_2\neq j_1}^{M}\sum_{j_1}^{M} P(K_{j_1} + K_{j_2} = n)$$

$$+ \sum_{j_3\neq j_1,j_2}^{M}\sum_{j_2\neq j_1}^{M}\sum_{j_1}^{M} P(K_{j_1} + K_{j_2} + K_{j_3} = n) + \cdots$$

$$+ \sum_{j_m\neq j_i,i=1,2,..,m}^{M} \ldots \sum_{j_2\neq j_1}^{M}\sum_{j_1}^{M} P(K_{j_1} + K_{j_2} + \ldots + K_{j_m} = n)$$

$$= e^{-\lambda}\frac{\lambda}{1!}\left(I_{\{k_1=n\}}\frac{\lambda_{k_1}}{\lambda}\right) + e^{-\lambda}\frac{\lambda^2}{2!}\left[I_{\{k_1+k_2=n\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)\right] + \cdots$$

$$+ e^{-\lambda}\frac{\lambda^3}{m!}\left[I_{\{k_1+k_2+\ldots+k_3=m\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}\times\ldots\times\lambda_{k_m}}{\lambda^3}\right)\right]$$

Penjelasan secara umum untuk $\left(e^{-\lambda}\frac{\lambda^m}{m!}\right)$ merupakan probabilitas untuk membentuk n dari m buah K (bagian Poissonnya), dan $\left(I_{\{k_1+k_2+\cdots+k_m=n\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}\times\ldots\times\lambda_{k_m}}{\lambda^m}\right)\right)$ merupakan probabilitas memilih m K mana yang digunakan untuk membentuk n, indikator $I_{\{k_1+k_2+\cdots+k_m=n\}}$ menyatakan m K tersebut yang dapat digunakan menghasilkan jumlah nilai n.

$$P(N=n)=e^{-\lambda}\frac{\lambda}{1!}\left(I_{\{k_1=n\}}\frac{\lambda_{k_1}}{\lambda}\right)+e^{-\lambda}\frac{\lambda^2}{2!}\left[I_{\{k_1+k_2=n\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}}{\lambda^2}\right)\right]+\cdots$$
$$+e^{-\lambda}\frac{\lambda^3}{m!}\left[I_{\{k_1+k_2+\ldots+k_3=m\}}\left(\frac{\lambda_{k_1}\times\lambda_{k_2}\times\ldots\times\lambda_{k_m}}{\lambda^3}\right)\right]$$
$$=\sum_{m=1}^{n}e^{-\lambda}\frac{\lambda^m}{m!}\sum_{k_1,\ldots,k_m\in\mathbb{N}^*}\mathrm{I}_{\{k_1+k_2+\cdots+k_m=n\}}\frac{\lambda_{k_1}\times\ldots\times\lambda_{k_m}}{\lambda^m}$$

dan

$$P(N=0)=e^{-\lambda}$$

Lampiran 2 Pembuktian Proposisi 2.1

Dengan menggunakan persamaan pada Lemma 1 dan definisi λ_k untuk distribusi *Geometric Poisson* yaitu

$$P(N=n) = \sum_{m=1}^{n} e^{-\lambda} \frac{\lambda^m}{m!} \sum_{k_1,\dots,k_m \in \mathbb{N}^*} \mathrm{I}_{\{k_1+k_2+\cdots+k_m=n\}} \frac{\lambda_{k_1} \times \dots \times \lambda_{k_m}}{\lambda^m}$$

$$\lambda_k = \lambda(1-\theta)^{k-1}\theta \quad \forall k \in \mathbb{N}^*$$

Pertama - tama perhatikan bentuk

$$\sum_{k_1,\dots,k_m \in \mathbb{N}^*} \frac{\lambda_{k_1} \times \dots \times \lambda_{k_m}}{\lambda^m}$$

Diketahui bahwa

$$\lambda_k = \lambda(1-\theta)^{k-1}\theta \quad \forall k \in \mathbb{N}^*$$

dimana

$\lambda_{k_1} = \lambda(1-\theta)^{k_1-1}\theta$, $\lambda_{k_2} = \lambda(1-\theta)^{k_2-1}\theta$, … , $\lambda_{k_m} = \lambda(1-\theta)^{k_m-1}\theta$,

maka

$$\sum_{k_1,\dots,k_m \in \mathbb{N}^*} \frac{\lambda_{k_1} \times \dots \times \lambda_{k_m}}{\lambda^m} = \frac{\lambda(1-\theta)^{k_1-1}\theta \times \lambda(1-\theta)^{k_2-1}\theta \times \dots \times \lambda(1-\theta)^{k_m-1}\theta}{\lambda^m}$$

$$= \frac{\lambda^m(1-\theta)^{(k_1+k_2+\cdots+k_m)-(1+1+\cdots+1)}\theta^m}{\lambda^m}$$

$$= (1-\theta)^{n-m}\theta^m,$$

didapat $\forall n \in \mathbb{N}^*$

$$P(N=n) = \sum_{m=1}^{n} e^{-\lambda} \frac{\lambda^m}{m!} (1-\theta)^{n-m} \theta^m \underbrace{\sum_{k_1,\ldots,k_m \in \mathbb{N}^*} \mathbb{I}_{\{k_1+\cdots+k_m=n\}}}_{A(n,k)}.$$

Lebih lanjut hanya tinggal dihitung nilai dari $A(n,k)$, yaitu banyaknya cara menghasilkan n dari penjumlahan k non negatif integer. Misalkan terdapat daftar 1,1,1 ... ,1 sebanyak n kali. Daftar ini memuat $n-1$ tempat untuk meletakkan tanda jumlah dan $k-1$ tanda jumlah. Sehingga jelas banyaknya cara menghasilkan n dari penjumlahan k non negatif integer adalah $A(n,k) = \binom{n-1}{k-1}$.

Sedemikian sehingga didapat

$$P(N=n) = \sum_{m=1}^{n} e^{-\lambda} \frac{\lambda^m}{m!} (1-\theta)^{n-m} \theta^m \binom{n-1}{m-1}$$

Lampiran 3 Pembuktian Proposisi 2.2

Berdasarkan transformasi *Kummer*, diketahui bahwa

$$M(a,b,z) = e^{z}M(b-a,b,-z)$$

maka untuk $e^{-z}M(n+1,2,z) = M(-n+1,2,-z)$ (Abramowitz & Stegun, 1972), berdasarkan definisi didapat

$$M(-n+1,2,-z) = 1 + \frac{(n-1)}{2}\frac{z}{1!} + \frac{(n-1)(n-2)}{2\times 3}\frac{z^2}{2!} + \cdots$$

$$= 1 + \sum_{m=2}^{\infty} \frac{(n-1)(n-2)\ldots(n-m+1)}{m!}\frac{z^{m-1}}{(m-1)!}$$

$$= 1 + \sum_{k=2}^{\infty} \frac{(n-1)(n-2)\ldots(n-m+1)}{(m-1)!}\frac{z^{m-1}}{m!}$$

$$= 1 + \sum_{m=2}^{\infty} \frac{(n-1)(n-2)\ldots(n-m+1)(n-m)!}{(m-1)!\,(n-m)!}\frac{z^{m-1}}{m!}$$

$$= 1 + \sum_{m=2}^{\infty} \frac{(n-1)!}{(m-1)!\,(n-1-(m-1))!}\frac{z^{m-1}}{m!}$$

$$= 1 + \sum_{m=2}^{\infty} \binom{n-1}{m-1}\frac{z^{m-1}}{m!}$$

$$= \binom{n-1}{1-1}\frac{z^{1-1}}{1!} + \sum_{m=2}^{\infty} \binom{n-1}{m-1}\frac{z^{m-1}}{m!}$$

$$= \sum_{m=1}^{\infty} \binom{n-1}{m-1}\frac{z^{m-1}}{m!}$$

$$= \sum_{m=1}^{n} \binom{n-1}{m-1} \frac{z^{m-1}}{m!} + \underbrace{\sum_{m=n+1}^{\infty} \binom{n-1}{m-1} \frac{z^{m-1}}{m!}}_{0}$$

$$= \sum_{m=1}^{n} \binom{n-1}{m-1} \frac{z^{m-1}}{m!}$$

Dengan

$$P(N=n) = \sum_{m=1}^{n} e^{-\lambda} \frac{\lambda^m}{m!} (1-\theta)^{n-m} \theta^m \binom{n-1}{m-1}$$

dengan nilai $z = \frac{\lambda\theta}{1-\theta}$

$$P(N=n) = \sum_{m=1}^{n} e^{-\lambda} (1-\theta)^n \left(\frac{\lambda\theta}{1-\theta}\right)^m \binom{n-1}{m-1} \frac{1}{m!}$$

$$= e^{-\lambda} (1-\theta)^n \sum_{m=1}^{n} z^m \binom{n-1}{m-1} \frac{1}{m!}$$

$$= e^{-\lambda} (1-\theta)^n z \sum_{m=1}^{n} \binom{n-1}{m-1} \frac{z^{m-1}}{m!}$$

$$= e^{-\lambda} (1-\theta)^n z M(-n+1, 2, -z)$$

$$= e^{-\lambda} (1-\theta)^n z e^{-z} M(n+1, 2, z)$$

Sehingga didapat $\forall n \in \mathbb{N}^*$ dengan $N \sim GP(\lambda, \theta)$ dengan $\lambda > 0$ dan $\theta \in [0,1]$

$$P(N=n) = e^{-\lambda} (1-\theta)^n z e^{-z} M(n+1, 2, z)$$

untuk M = *Confluent Hypergeometric Function*, dan $z = \frac{\lambda\theta}{1-\theta}$.