

# Pengaruh Tingkat Pendidikan, Tingkat Kesejahteraan dan Penghasilan Terhadap Konsumsi Rokok Harian dari Penduduk Dewasa di Indonesia Tahun 2015

Pardomuan Robinson Sihombing<sup>1</sup>

Ade Marsinta Arsani<sup>2</sup>

Badan Pusat Statistik Republik Indonesia

## Abstraksi

Saat ini, Indonesia menjadi negara ketiga dengan jumlah perokok tertinggi di dunia setelah Cina dan India. Kerugian makro ekonomi akibat konsumsi rokok di Indonesia pada 2015 mencapai hampir Rp 600 triliun. Ada banyak faktor yang dapat mengakibatkan seseorang mengkonsumsi rokok di antara dari segi sosio ekonomi, demografi, lingkungan, budaya dan lainnya. Mengingat konsumsi rokok yang tinggi di Indonesia serta faktor risiko yang terjadi akibat mengkonsumsi rokok, maka penelitian ini ingin mengetahui faktor-faktor yang mempengaruhi jumlah batang rokok yang dihisap. Jumlah rokok yang dihisap setiap hari merupakan data cacah nonnegatif. Untuk pemodelan variabel respon yang berupa data cacah, model yang biasa digunakan adalah regresi Poisson, regresi Binomial, dan regresi Negative Binomial. Konsumsi rokok dalam batang per hari merupakan salah satu kasus data cacahan dengan banyak nilai 0 (excess zero). Untuk mengatasi masalah overdispersion yang terjadi, salah satu cara adalah menggunakan Zero Inflated Negative Binomial (ZINB) atau Hurdle Negative Binomial (HNB). Kedua model tersebut digunakan untuk memodelkan data count dengan banyak nilai 0 pada respon dan terjadi overdispersion. Data konsumsi rokok yang dihasilkan dari IFLS memiliki nilai zero excess dan terdapat overdispersi. Model ZINB lebih baik daripada model HNP karena memiliki nilai AIC dan BIC yang lebih kecil. Pada model log hanya variabel penghasilan yang mempengaruhi peluang mengkonsumsi merokok. pada model logit hanya variabel dummy SMP yang tidak mempengaruhi peluang untuk tidak mengkonsumsi rokok, sedangkan variabel lainnya pendidikan, kesejahteraan dan penghasilan mempengaruhi peluang tidak mengkonsumsi rokok. Semakin tinggi pendidikan dan kesejahteraan akan meningkatkan peluang orang untuk tidak merokok

**Kata Kunci:** rokok; overdispersi; binomial negatif.

---

<sup>1</sup> Pardomuan Robinson adalah Statisti Ahli Pertama di Badan Pusat Statistik Republik Indonesia.

<sup>2</sup> Ade Marsinta adalah mahasiswa di Universitas Padjajaran dan dan Statistisi di Direktorat Neraca Pengeluaran, Badan Pusat Statistik Republik Indonesia. E-mail: ademarsinta@gmail.com

## **Pengaruh Tingkat Pendidikan, Tingkat Kesejahteraan dan Penghasilan Terhadap Konsumsi Rokok Harian dari Penduduk Dewasa di Indonesia Tahun 2015**

Ade Marsinta Arsani dan Pardomuan Robinson Sihombing

### **I. Pendahuluan**

#### **1.1. Latar Belakang**

Merokok merupakan salah satu kebiasaan berisiko tinggi. Badan Kesehatan Dunia atau World Health Organization (WHO) melaporkan kebiasaan atau perilaku merokok merupakan salah satu penyebab kematian paling besar di dunia. Dalam laporan WHO dilaporkan bahwa kematian dini yang diakibatkan oleh penyakit yang terkait dengan kebiasaan dan perilaku merokok seperti kanker, penyakit jantung, penyakit hati dan stroke mencapai 5 juta jiwa lebih setiap tahun (WHO, 2003). Hal ini karena rokok memiliki faktor risiko berbagai macam penyakit, antara lain penyakit kardiovaskular, serebrovaskular, impotensi, berbagai jenis kanker (Alit, dkk dalam Santi, 2013). Merokok menjadi kebiasaan yang merugikan kesehatan dimana kebiasaan ini sering kali sulit dihentikan karena adanya efek ketergantungan yang ditimbulkan oleh nikotin yang terkandung di dalamnya. Selain itu akibat yang ditimbulkan yaitu berupa penyakit akibat rokok terjadi dalam jangka waktu yang cukup lama, sehingga sering kali menyebabkan kegagalan dalam upaya mencegah untuk tidak merokok atau menghentikan kebiasaan merokok (Dachroni, 1990). Kebiasaan merokok ini tidak hanya berdampak pada perokok itu sendiri akan tetapi juga bagi perokok pasif, terutama pada kelompok rentan seperti usia balita, anak sekolah, dan populasi perempuan meskipun ada sebagian dari perempuan juga merokok. Perilaku merokok itu sendiri dapat diukur salah satunya melalui banyaknya jumlah rokok yang dihisap (Fitriani, 2012, dalam Hestiana, 2013).

Indonesia merupakan negara keempat dengan konsumsi rokok terbesar di dunia setelah China, USA, dan Rusia. Jumlah batang rokok yang dikonsumsi di Indonesia cenderung meningkat dari 182 milyar batang pada tahun 2001 menjadi 260,8 milyar batang pada tahun 2009. Angka prevalensi rokok pada tahun 2010 pun cukup tinggi yaitu 34,7 persen (Buku Fakta Tentang Tembakau, 2012). Saat ini, Indonesia menjadi negara ketiga dengan jumlah perokok tertinggi di dunia setelah Cina dan India. Kerugian makro ekonomi akibat konsumsi rokok di Indonesia pada 2015 mencapai hampir Rp 600 triliun, atau empat kali lipat dari jumlah cukai rokok pada tahun yang sama (sains.kompas.com). Ada banyak faktor yang dapat mengakibatkan seseorang mengkonsumsi rokok di antara dari segi sosio ekonomi, demografi, lingkungan, budaya dan lainnya. Jika dikaitkan dengan kondisi sosio ekonomi dari perokok, beberapa studi menemukan bahwa perilaku merokok berhubungan terbalik dengan tingkat pendidikan, level pendapatan, dan jenis pekerjaan. Studi lain mencoba mengaitkan kondisi kemiskinan dan perilaku merokok. Perokok aktif ditemukan berkorelasi dengan persentil pendapatan paling rendah (Azizah dkk, 2013). Berdasarkan hasil studi dari Nutrition & Health Surveillance System tahun 1998 – 2000, secara umum menyimpulkan bahwa perilaku merokok dapat memperburuk tingkat kemiskinan, karena pendapatan keluarga yang terbatas malah diprioritaskan untuk belanja tembakau dan produk tembakau. Pada kelompok pendapatan kuintil termiskin, dilaporkan bahwa belanja untuk tembakau ternyata 3 (tiga) kali dari belanja pendidikan anak, dan sebesar 4,3 kali biaya kesehatan keluarga.

Analisis regresi umumnya digunakan untuk mengetahui hubungan suatu variabel respon dengan satu atau beberapa variabel prediktor. Jumlah rokok yang dihisap setiap hari merupakan data cacah nonnegatif. Untuk pemodelan variabel respon yang berupa data cacah, model yang biasa digunakan adalah regresi Poisson, regresi Binomial, dan regresi Negative Binomial (Agresti, 2002). Konsumsi rokok dalam batang per hari merupakan salah satu kasus

data cacahan dengan banyak nilai 0 (*excess zero*). Jumlah angka nol besar akan menyebabkan sebaran data tidak sesuai dengan distribusi standar. Sebagaimana metode regresi Poisson akan menjadi tidak sesuai jika banyak data bernilai nol yang mengakibatkan terjadinya *overdispersion* (Ridout, Demetrio, dan Hindie, 1998). Model yang biasanya digunakan pada kasus *excess zero* adalah regresi Zero Inflated Poisson (ZIP) atau Hurdle Poisson (HP). Permasalahan yang sering muncul pada kedua pemodelan tersebut adalah terjadinya *overdispersion*. Untuk mengatasi masalah *overdispersion* yang terjadi, salah satu cara adalah menggunakan Zero Inflated Negative Binomial (ZINB) atau Hurdle Negative Binomial (HNB). Kedua model tersebut digunakan untuk memodelkan data count dengan banyak nilai 0 pada respon dan terjadi *overdispersion*. Pemodelan dengan Zero Inflated merupakan model mixture dengan memodelkan observasi bernilai 0 dengan proses biner dan model independen Poisson/ Negatif Binomial untuk observasi bernilai positif. Model Hurdle diasumsikan sebagai dua proses independen fungsi dari variabel respon yaitu model biner untuk observasi bernilai nol dan Truncated Poisson/Negatif Binomial untuk observasi bernilai positif. Kedua model ini sama-sama menggunakan metode maksimum likelihood untuk mendapatkan estimasi parameter yang digunakan.

## 1.2. Rumusan Masalah

Mengingat konsumsi rokok yang tinggi di Indonesia serta faktor risiko yang terjadi akibat mengkonsumsi rokok, maka penelitian ini ingin mengetahui faktor-faktor yang mempengaruhi jumlah batang rokok yang dihisap. Variabel independen yang mempengaruhi konsumsi jumlah batang rokok yang dihisap dibatasi hanya 3 variabel yaitu tingkat pendidikan, tingkat kesejahteraan dan penghasilan per bulan. Dalam penelitian ini, konsumsi rokok per hari individu akan dimodelkan dengan Zero Inflated Poisson (ZIP), Zero Inflated Negative Binomial (ZINB), Hurdle Poisson (HP) dan Hurdle Negative Binomial (HNB). Metode/alat analisis tersebut sesuai untuk memodelkan data cacah serta dengan kondisi *excess zero*, yaitu kondisi dimana data respon banyak bernilai nol, dan dapat menanggapi *overdispersi* yang umumnya sering muncul pada analisis data cacah.

Berdasarkan uraian tersebut, permasalahan dalam penelitian ini adalah sebagai berikut:

1. Bagaimana karakteristik masyarakat Indonesia yang mengonsumsi rokok?
2. Faktor-faktor apa yang memengaruhi konsumsi rokok di Indonesia?
3. Bagaimana perbandingan metode Zero Inflated maupun Hurdle pada distribusi Poisson dan Negatif Binomial dalam memodelkan faktor apa yang memengaruhi konsumsi rokok.

## II. Tinjauan Pustaka

### 2.1. Penelitian Terdahulu

Pada umumnya perokok dapat dibagi atas tiga kategori, yaitu : 1) bukan perokok (*non smokers*), adalah seseorang yang belum pernah mencoba merokok sama sekali; 2) perokok eksperimen (*experimental smokers*), adalah seseorang yang telah mencoba merokok tetapi tidak menjadikannya sebagai suatu kebiasaan; dan 3) perokok tetap atau perokok reguler (*regular smokers*), adalah seseorang yang teratur merokok baik dalam hitungan mingguan atau dengan intensitas yang lebih tinggi lagi.

Berdasarkan kemampuannya menghisap rokok dalam sehari, perokok dapat dibagi menjadi tiga kelompok, yaitu:

1. Perokok ringan, yaitu perokok yang mengonsumsi rokok hingga 10 batang per hari.
2. Perokok sedang, yaitu perokok yang mengonsumsi 11-20 batang rokok per hari.
3. Perokok berat, yaitu perokok yang mengonsumsi lebih dari 20 batang rokok per hari.

Perilaku merokok seseorang dapat kita identifikasi berdasarkan karakteristik individunya. Sugiharti, Sukartini, dan Handriana melalui penelitiannya menyebutkan bahwa,

individu dengan pendidikan setara sekolah dasar (SD) mempunyai kecenderungan merokok lebih besar dibanding individu dengan jenjang pendidikan yang lebih tinggi. Artinya, semakin rendah tingkat pendidikan seseorang, maka semakin besar peluang orang tersebut menjadi perokok.

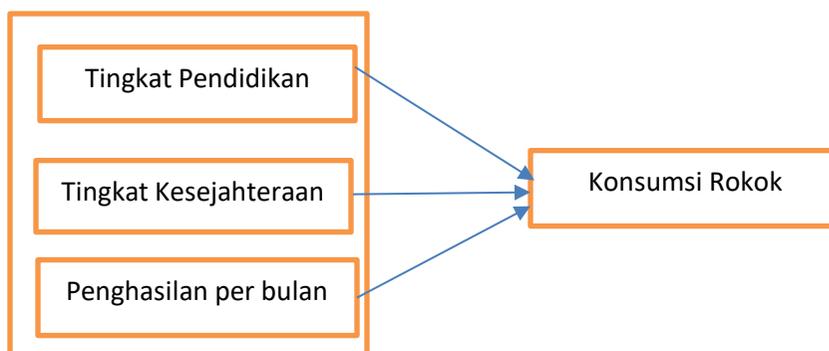
Penelitian di Amerika Serikat menjelaskan adanya hubungan yang erat antara kebiasaan merokok dan latar pendidikan sang perokok. *Centers for Disease Control and Prevention* (CDC) dalam laporannya yang berjudul *Morbidity and Mortality Weekly Report*, 2008 menyatakan perokok dengan pendidikan diploma menduduki peringkat tertinggi yaitu sebesar 44%. Sementara itu, perokok yang pernah mengenyam pendidikan 9-11 tahun mempunyai tingkat prevalensi 33.3%, dan perokok yang berlatar pendidikan perguruan tinggi hanya 11.4%. Prevalensi perokok berpendidikan sarjana jauh lebih rendah lagi, yaitu hanya 6.2%.

Kesejahteraan konsumen akan menentukan besarnya daya beli yang dimilikinya. Pada barang normal, peningkatan kesejahteraan dan pendapatan konsumen akan menaikkan permintaan barang tersebut. Sebaliknya untuk barang inferior, peningkatan kesejahteraan konsumen justru akan menurunkan konsumsinya. Barang yang digunakan orang kaya akan berkurang permintaannya, namun orang kurang sejahtera yang mengalami peningkatan pendapatan akan bertambah permintaannya (Sukirno 2008). Sebagian besar komoditi rokok merupakan barang normal di mana kenaikan kesejahteraan dan pendapatan akan meningkatkan demand untuk komoditi tersebut. Pada masyarakat miskin, kenaikan pendapatan akan meningkatkan konsumsi rokok sebesar 6% sementara untuk orang kaya hanya 2.1%.

Dari sisi pendapatan yang diproksi dengan pengeluaran, semakin tinggi level pengeluaran seseorang sampai pada persentil 3, semakin besar peluang seseorang menjadi perokok. Ahsan (2006) melalui penelitiannya menyebutkan bahwa salah satu variabel yang signifikan mempengaruhi probabilitas seseorang menjadi perokok tingkat pendapatan.

## 2.2. Kerangka Pemikiran

Berdasarkan kajian literatur dan teori yang ada maka dapat dibentuk kerangka pemikiran sebagai berikut:



## 2.3. Hipotesis Penelitian

Berdasarkan tinjauan pustaka, teori yang relevan dan kerangka penelitian di atas maka hipotesis dalam penelitian ini adalah:

- Terdapat pengaruh tingkat pendidikan terhadap konsumsi rokok di Indonesia
- Terdapat pengaruh tingkat kesejahteraan terhadap konsumsi rokok di Indonesia
- Terdapat pengaruh penghasilan per bulan terhadap konsumsi rokok di Indonesia

## III. Metode Penelitian

**3.1. Data Penelitian**

Penelitian ini menggunakan data sekunder yang berasal dari *Survey Indonesia Family Life Survey* Tahun 2015 ([www.rand.org](http://www.rand.org)). Berdasarkan hasil penelitian sebelumnya, maka variabel yang digunakan pada penelitian ini adalah konsumsi batang rokok/hari sebagai variabel respon/dependen/terikat (Y). Adapun yang menjadi variabel prediktor/independen/bebas adalah Pendidikan (X1), tingkat kesejahteraan (X2) dan penghasilan/bulan (X3). Data akan dimodelkan menggunakan regresi *Zero Inflated* dan *Hurdle* baik untuk distribusi Poisson maupun Negatif Binomial. Berikut disajikan lebih rinci variabel penelitian yang digunakan:

**Tabel 1.** Variabel Penelitian

Variabel	Deskripsi	Keterangan	Skala Data
Y	Jumlah Konsumsi Rokok/hari	dalam satuan batang	Rasio
X1	Pendidikan	1=sd 2=smp 3=sma 4>sma	Ordinal
X2	Tingkat Kesejahteraan	1=sangat miskin 2=miskin 3=cukup miskin 4=cukup kaya 5=kaya 6=sangat kaya	Ordinal
X3	Penghasilan/bulan	dalam rupiah	Rasio

Pada variabel X1 (tingkat pendidikan) dan X2 (tingkat kesejahteraan) merupakan data kategorik yang berskala ordinal sehingga dalam analisis diperlakukan sebagai variabel *dummy*. Banyaknya variabel *dummy* yang terbentuk adalah sebanyak jumlah kategori dikurangi 1. Untuk variabel tingkat pendidikan karena terdapat 4 kategori maka ada sebanyak 3 variabel *dummy* dengan kategori 1 (SD) sebagai kategori referensi sedangkan pada variabel tingkat kesejahteraan terdapat 6 kategori sehingga terdapat 5 variabel *dummy* dengan kategori 1 (sangat miskin) sebagai variabel referensi.

**3.2. Metodologi**

Data diskrit adalah data *count*, yaitu data yang nilainya non-negatif dan menyatakan banyaknya kejadian dalam interval waktu, ruang, atau volume tertentu. Beberapa distribusi yang mempunyai bentuk fungsi peluang dan nama tertentu dari peubah acak diskrit, yaitu:

**3.2.1. Zero-Inflated Negative Binomial (ZINB)**

Berdasarkan analisis yang dilakukan dengan ZIP atau *Hurdle* dapat diketahui apakah asumsi *mean* sama dengan variansinya atau tidak. Jika ternyata diketahui rata-ratanya tidak sama dengan variansinya (overdispersi), maka masalah ini dapat ditanggulangi dengan menerapkan model ZINB. Pada model ini data cacahan diasumsikan diperoleh dari dua proses, yaitu proses pertama merupakan munculnya data *count* yang hanya berisi nilai nol ditentukan dengan probabilitas *p*. Proses kedua yaitu nilai nol dan nilai positif pada data cacahan, kedua-duanya dihasilkan oleh suatu proses yang mengikuti distribusi *Negative Binomial* (NB). Model ZINB diperoleh dari persamaan:

$$F_i = \frac{\exp(\mathbf{z}_i\gamma)}{1 + \exp(\mathbf{z}_i\gamma)} \quad (1)$$

$$\text{logit}(F_i) = \log\left(\frac{F_i}{1-F_i}\right) = \mathbf{z}_i\gamma; (i = 1, 2, \dots, n) \quad (2)$$

Dengan  $g(y)$  adalah fungsi distribusi NB, sehingga dihasilkan model ZINB sebagai berikut:

$$\Pr(Y_i = y_i | \mathbf{x}_i, \mathbf{z}_i) = \begin{cases} F_i + (1 - F_i)(1 + \alpha\mu_i)^{-\alpha^{-1}} & ; y_i = 0 \\ (1 - F_i) \frac{\Gamma(y_i + \alpha^{-1})}{y_i! \Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right)^{\alpha^{-1}} \left( \frac{\mu_i}{\alpha^{-1} + \mu_i} \right)^{y_i} & ; y_i > 0 \end{cases} \quad (3)$$

maka model ZINB dapat ditulis sebagai berikut:

$$\Pr(Y_i = y_i | \mathbf{x}_i, \mathbf{z}_i) = \begin{cases} \frac{\exp(\mathbf{z}_i \gamma)}{1 + \exp(\mathbf{z}_i \gamma)} + \frac{1}{1 + \exp(\mathbf{z}_i \gamma)} (1 + \alpha \exp(\mathbf{x}_i \beta))^{-\alpha^{-1}} & ; y_i = 0 \\ \frac{1}{1 + \exp(\mathbf{z}_i \gamma)} \frac{\Gamma(y_i + \alpha^{-1})}{y_i! \Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \exp(\mathbf{x}_i \beta)} \right)^{-\alpha} & ; y_i > 0 \end{cases} \quad (4)$$

### 3.2.2. Model Hurdle Negative Binomial (HNB)

Menurut Pontoh (2015) model *Hurdle Negative Binomial* merupakan alternatif untuk data yang memiliki banyak nilai nol dan mengatasi terjadinya overdispersi. Misalkan bahwa  $Y_i$  merupakan variabel random diskrit dengan  $i$  adalah bilangan bulat non negatif ( $i = 1, 2, \dots, n$ ) dan  $Y_i$  merupakan variabel respon dari model regresi HNB, maka nilai dari variabel respon, dimodelkan dalam dua keadaan. Keadaan pertama disebut zero state dan menghasilkan hanya pengamatan bernilai nol, sementara keadaan kedua disebut *negative binomial state* yang memiliki sebaran Binomial Negatif

### 3.3. Prosedur Penelitian

Langkah-langkah yang dilakukan dalam penelitian:

1. Melakukan analisis deskriptif variabel penelitian.
2. Melakukan pengecekan excess zero dengan histogram dan tabel frekuensi
3. Melakukan analisis dan penaksiran parameter dengan Regresi Zero Inflated Poisson (ZIP) dan Hurdle Poisson (HP)
4. Melakukan pengecekan overdispersi
5. Melakukan analisis dan penaksiran parameter dengan Regresi Zero Inflated Negative Binomial (ZINB) dan Hurdle Negative Binomial (HNB)
6. Pemilihan model terbaik antara ZINB dan HNB dengan metode AIC dan BIC
7. Melakukan pengujian kelayakan model
8. Interpretasi hasil

## IV. Hasil dan Pembahasan

### 4.1. Analisis Deskriptif

Tingkat Pendidikan

**Tabel 2.** Tingkat Pendidikan Responden

Pendidikan	Frekuensi	Percent
SD	863	13.9
SMP	1081	17.4
SMA	2469	39.7
> SMA atau PT	1800	29
Total	6213	100

Sumber: data diolah

Dari hasil di atas terlihat bahwa responden kebanyakan berada pada tingkat Pendidikan SMA sebesar 2469 responden atau sekitar 39.7 persen.

Tingkat Kesejahteraan

**Tabel 3.** Tingkat Pendidikan Responden

Kesejahteraan	Frekuensi	Percent
sangat miskin	324	5.2
miskin	995	16.0
cukup miskin	3066	49.3
cukup kaya	1639	26.4
kaya	138	2.2
sangat kaya	51	0.8
Total	6213	100.0

**Sumber:** data diolah

Penghasilan per bulan

**Tabel 4.** Tingkat Penghasilan Responden Per Bulan

Variabel	N	Minimum	Maximum	Mean	Std. Deviation
penghasilan/bulan	6213	2500	99999997	2034140.7	12994469.8

**Sumber:** data diolah

Dari hasil di atas dapat diketahui bahwa nilai minimum penghasilan responden per bulan sebesar 2500 sedangkan nilai maksimum sebesar 99999997 dan nilai standar deviasi lebih besar dari nilai rata-ratanya. Sehingga untuk analisis inferensia maka variabel penghasilan per bulan di transformasi dengan logaritma natural.

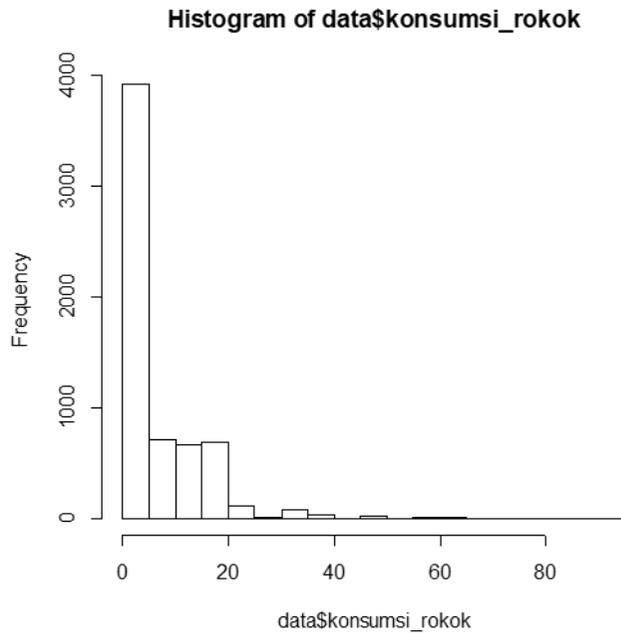
Konsumsi Rokok per hari

**Tabel 5.** Konsumsi Rokok Per Hari

Variabel	N	Minimum	Maximum	Median	Mean	Std. Deviation	Variance
konsumsi batang rokok/hari	6213	.00	95.00	.00	5.5492	8.24401	67.964

**Sumber:** data diolah

Dari hasil di atas dapat diketahui bahwa nilai minimum konsumsi rokok responden per hari adalah 0, begitu pun nilai mediannya adalah 0. Nilai mean lebih kecil dari nilai varian data.



**Gambar 1.** Frekuensi Konsumsi Rokok

Untuk memudahkan dalam intrepetasi data konsumsi rokok, dibuat tabel deskriptif kategori perokok sebagai berikut:

**Tabel 6.** Tingkat Konsumsi Rokok

Tingkat Konsumsi Rokok	Frekuensi	Percent
tidak merokok (0 batang)	3400	54.7
perokok ringan (1-10 batang)	1225	19.7
perokok sedang (11-20 batang)	1332	21.4
perokok berat (>30 batang)	256	4.1
Total	6213	100.0

**Sumber:** data diolah

Berdasarkan hasil data di atas dapat dilihat bahwa dari 6213 responden terdapat 3400 responden atau sekitar 54,4 persen responden yang tidak merokok, dengan kata lain terjadi *excess zero value* di dalam data. Berdasarkan fenomena di atas maka model awal yang digunakan adalah *Zero Inflated Poisson* (ZIP) dan *Hurdle Poisson* (HP) dan akan dicek overdispersi model.

Selanjutnya akan dilihat bagaimana konsumsi rokok per tingkatan pendidikan dan tingkatan kesejahteraan.

**Tabel 7.** Hubungan Antara Tingkat Pendidikan dan Konsumsi Rokok

Tingkat Pendidikan	Konsumsi Rokok/hari	
	Mean	SD
SD	7.24	8.16
SMP	6.73	8.31
SMA	5.89	8.22
>SMA atau PT	3.56	7.89

Sumber: data diolah

Dari tabel di atas dapat dilihat bahwa semakin tinggi tingkat pendidikan maka konsumsi rokok/hari semakin sedikit. Akan tetapi variasi konsumsi rokok terbesar terdapat pada tingkatan pendidikan SMP.

**Tabel 8.** Hubungan Antara Tingkat Kesejahteraan dan Konsumsi Rokok

Tingkat Kesejahteraan	Konsumsi Rokok/hari	
	Mean	SD
Sangat Miskin	8.42	8.2
Miskin	6.95	7.93
Cukup Miskin	5.66	8.38
Cukup Kaya	3.95	7.72
Kaya	4.34	7.94
Sangat Kaya	7.96	12.1

Sumber: data diolah

Dari tabel di atas dapat dilihat bahwa tingkat konsumsi rokok terendah pada tingkat kesejahteraan menengah (cukup miskin dan cukup kaya). Konsumsi rokok tertinggi pada tingkatan kesejahteraan sangat miskin dan sangat kaya.

**A. Pemodelan Regresi *Zero Inflated Negatif Binomial* dan *Hurdle Negative Binomial***

Karena model ZIP dan HP terjadi overdispersi, maka model yang lebih tepat digunakan adalah ZINB dan HNB. Sama dengan model ZIP dan Hurdle Poisson, kedua model inipun terdiri dari dua model, yaitu model Log dan Model Logit.

**Tabel 9.** Model ZINB dan HNB

Parameter	Model ZINB				Model HNB			
	Estimasi	Standard Error	p-values	ket	Estimasi	Standard Error	p-values	ket
Model log								
(Intercept)	0.906	0.181	0.000	***	0.910	0.181	0.000	***
pendidikan SMP	-0.022	0.039	0.565		-0.022	0.039	0.567	
pendidikan SMA	-0.042	0.035	0.230		-0.042	0.035	0.230	
pendidikan >SMA	0.002	0.043	0.961		0.002	0.043	0.963	

Parameter	Model ZINB				Model HNB			
	Estimasi	Standard Error	p-values	ket	Estimasi	Standard Error	p-values	ket
kesejahteraan miskin	-0.056	0.051	0.271		-0.056	0.051	0.272	
Kesejahteraan cukup miskin	-0.035	0.048	0.460		-0.035	0.048	0.461	
Kesejahteraan cukup kaya	-0.077	0.053	0.147		-0.077	0.053	0.147	
kesejahteraan kaya	0.001	0.104	0.990		0.001	0.104	0.995	
Kesejahteraan sangat kaya	0.234	0.135	0.085		0.234	0.135	0.084	
log(penghasilan)	0.117	0.013	0.000	***	0.117	0.013	0.000	***
Model logit (Intercept)	3.834	0.380	0.000	***	-3.892	0.377	0.000	***
pendidikan SMP	0.115	0.097	0.238		-0.114	0.096	0.234	
pendidikan SMA	0.423	0.086	0.000	***	-0.420	0.085	0.000	***
pendidikan >SMA	1.343	0.096	0.000	***	-1.335	0.095	0.000	***
kesejahteraan miskin	0.425	0.143	0.003	**	-0.422	0.140	0.003	**
Kesejahteraan cukup miskin	0.880	0.133	0.000	***	-0.871	0.131	0.000	***
Kesejahteraan cukup kaya	1.360	0.140	0.000	***	-1.350	0.138	0.000	***
kesejahteraan kaya	1.333	0.229	0.000	***	-1.321	0.227	0.000	***
Kesejahteraan sangat kaya	0.949	0.320	0.003	**	-0.934	0.318	0.003	**
log(penghasilan)	-0.367	0.027	0.000	***	0.369	0.027	0.000	***
AIC	26587.95				26587.970			
BIC	26729.37				26729.390			

Sumber: data diolah

Keterangan: \*\*\*signifikan 0.1%, \*\* signifikan pada 1 %, \* signifikan pada 5 %

Dari hasil di atas di dapat bawa, nilai AIC dan BIC model ZINB lebih kecil daripada model HNB maka model yang digunakan adalah model ZINB. Sedangkan dari hasil uji Vuong Non dengan membandingkan model ZIP dan ZINB di dapat z stat 13.75146 dan  $p.value < 2.22e-16$  sehingga dapat disimpulkan ada perbaikan dengan menggunakan model ZINB dibandingkan model ZIP.

### B. Pemodelan dan Pengujian Parameter Regresi Zero-Inflated Negative Binomial (ZINB)

Pemodelan regresi ZINB dengan memasukkan semua variabel bebas menghasilkan estimasi parameter seperti terlihat pada tabel 4.10. Selanjutnya untuk menguji kelayakan model ZINB dilakukan pengujian secara serentak (*overall*), dengan hipotesis:

Ho: Semua variabel prediktor pada model tidak mempengaruhi jumlah konsumsi rokok

$$(\hat{\beta}_1 = \hat{\beta}_2 = \dots = \hat{\beta}_j \text{ atau } \hat{\alpha}_1 = \hat{\alpha}_2 = \dots = \hat{\alpha}_j = 0 \text{ dimana } j \text{ adalah banyaknya variabel prediktor})$$

H1: Paling sedikit ada satu variabel prediktor pada model yang mempengaruhi jumlah konsumsi rokok (paling sedikit ada satu  $\hat{\beta}_j \neq 0$  atau  $\hat{\alpha}_i \neq 0$  di mana  $j = 1, 2, \dots, k$ )

Statistik yang digunakan untuk pengujian secara serentak adalah dengan menggunakan LRT. Pada tabel, terlihat bahwa nilai statistik GHitung ( $> \chi^2_{0.05, 21} = 25.00$ ). Kesimpulannya adalah menolak H0 yang berarti bahwa minimal ada satu variabel prediktor yang berpengaruh terhadap variabel respon.

Selanjutnya dilakukan pengujian secara parsial untuk mengetahui efek yang diberikan masing-masing variabel bebas. Pengujian parameter parsial untuk model log digunakan hipotesis sebagai berikut:

$$Ho: \hat{\beta}_j = 0$$

$$H1: \hat{\beta}_j \neq 0$$

Statistik uji yang digunakan adalah Uji Wald yaitu  $W_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$ ;  $W_j = \frac{\hat{\gamma}_j}{SE(\hat{\gamma}_j)}$ . Pada tabel 3, diketahui bahwa pada taraf signifikansi 5 persen, estimasi parameter yang signifikan untuk model log adalah koefisien  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$  dan  $\hat{\beta}_5$ . Sedangkan hipotesis untuk pengujian parameter secara parsial untuk model logit adalah sebagai berikut:

$$H_0: \hat{\alpha}_j = 0$$

$$H_1: \hat{\alpha}_j \neq 0$$

Pada taraf 5 persen diperoleh parameter yang signifikan adalah  $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4, \hat{\alpha}_5$ .

Untuk model log ZINB hanya variabel penghasilan yang signifikan, sedangkan untuk model logit ZINB hanya variabel dummy pendidikan SMP yang tidak signifikan.

### C. Interpretasi Model Regresi ZINB

Model yang dihasilkan pada regresi ZINB (tabel 4.10) terdiri dari dua bagian, yaitu model dengan bentuk log dan logit. Bentuk log menyatakan peluang variabel respon Yi bernilai positif dipengaruhi oleh variabel prediktor yang signifikan, sedangkan bentuk logit menyatakan peluang respon Yi bernilai nol dipengaruhi oleh variabel-variabel bebas yang signifikan.

Model log memperlihatkan beberapa hal, diantaranya adalah peningkatan penghasilan 1 persen akan cenderung meningkatkan mengkonsumsi rokok sebesar  $\exp(0.117) = 1.1246$ . Sedangkan variabel pendidikan dan kesejahteraan belum menunjukkan pengaruh yang signifikan terhadap konsumsi rokok.

Model logit menjelaskan peluang variabel respon bernilai nol dipengaruhi oleh variabel prediktor yang signifikan. Belum cukup bukti menyatakan bahwa terdapat perbedaan peluang tidak mengkonsumsi rokok antara tidak pendidikan SMP dengan Pendidikan SD. Tingkat pendidikan SMA, peluang untuk tidak mengkonsumsi rokok  $\exp(0.423) = 1.52$  kali lebih besar dibanding penduduk yang berstatus Pendidikan SD. Tingkat pendidikan lebih tinggi dari SMA atau PT, peluang untuk tidak mengkonsumsi rokok  $\exp(1.343) = 3.83$  kali lebih besar dibanding penduduk yang berstatus Pendidikan SD. Secara umum dapat disimpulkan bahwa semakin tinggi pendidikan maka peluang untuk tidak mengkonsumsi rokok semakin tinggi. Hal ini mengindikasikan bahwa semakin tinggi pendidikan, maka pengetahuan tentang bahaya rokok umumnya akan semakin tinggi. Hal ini senada dengan hasil penelitian yang dilakukan oleh Sirait, Pradono, & Toruan (2002) yang menemukan bahwa program pendidikan kesehatan di sekolah menjadi sangat penting karena perokok biasanya mulai merokok sejak di bangku sekolah. Pada murid diajarkan bukan hanya tentang bahaya merokok akan tetapi juga cara menolak tawaran merokok.

Variabel lain yang juga signifikan mempengaruhi kecenderungan merokok adalah tingkat kesejahteraan. Pada kelompok tingkat kesejahteraan miskin, peluang untuk tidak mengkonsumsi rokok  $\exp(0.425) = 1.52$  kali lebih besar dibanding penduduk yang berstatus sangat miskin. Sementara itu, pada kelompok tingkat kesejahteraan cukup miskin, peluang untuk tidak mengkonsumsi rokok  $\exp(0.880) = 2.41$  kali lebih besar dibanding penduduk yang berstatus sangat miskin. Pada kelompok tingkat kesejahteraan cukup kaya, peluang untuk tidak mengkonsumsi rokok  $\exp(1.360) = 3.89$  kali lebih besar dibanding penduduk yang berstatus sangat miskin. Selanjutnya, kelompok tingkat kesejahteraan kaya mempunyai peluang untuk tidak mengkonsumsi rokok  $\exp(1.333) = 3.79$  kali lebih besar dibanding penduduk yang berstatus sangat miskin. Kecenderungan ini melemah pada kelompok tingkat kesejahteraan sangat kaya, peluang untuk tidak mengkonsumsi rokok adalah  $\exp(0.949)$  atau hanya 2.58 kali lebih besar dibanding penduduk yang berstatus sangat miskin. Lima point di atas mengindikasikan bahwa secara umum, meskipun tidak linier, semakin tinggi tingkat kesejahteraan, maka peluang untuk mengkonsumsi rokok akan semakin kecil. Hasil senada juga ditemukan dalam penelitian yang dilakukan oleh Hamzah (2012) dan Afif & Sasana (2019) yang menyatakan bahwa masyarakat miskin yang cenderung lebih banyak

mengonsumsi rokok. Hasil pada variable kategorik ini juga selaras dengan variabel kuantitatif penghasilan yang menunjukkan bahwa ketika terdapat kenaikan tingkat penghasilan maka peluang untuk mengonsumsi rokok akan turun.

## V. Kesimpulan dan Saran

### 5.1. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang sudah diuraikan, dapat disimpulkan bahwa;

1. Data konsumsi rokok yang dihasilkan dari IFLS memiliki nilai *zero excess* dan terdapat overdispersi
2. Model ZINB lebih baik daripada model HNP karena memiliki nilai AIC dan BIC yang lebih kecil
3. Model ZINB memperbaiki model ZIP yang mengalami overdispersi
4. Pada model log ZINB hanya variabel penghasilan yang mempengaruhi peluang mengonsumsi merokok
5. Pada model logit ZINB hanya variabel dummy SMP yang tidak mempengaruhi peluang untuk mengonsumsi rokok, sedangkan variabel lainnya pendidikan, kesejahteraan dan penghasilan mempengaruhi peluang mengonsumsi rokok
6. Semakin tinggi pendidikan dan kesejahteraan akan meningkatkan peluang orang untuk tidak merokok

### 5.2. Saran

Adapun saran yang dapat dilakukan dalam penelitian selanjutnya adalah dengan membuat variabel penghasilan/pendapatan menjadi variabel kategorik berskala ordinal serta menambahkan variabel independen lainnya yang turut mempengaruhi seseorang untuk mengonsumsi rokok baik secara ekonomi (seperti harga rokok, barang substitusi dan komplementer) maupun sosial demografi (seperti jenis kelamin, jumlah anggota keluarga, dan lainnya).

## Referensi

- Ahsan A. 2006. Pengaruh Faktor Sosial Ekonomi terhadap Perilaku Merokok Individu: Analisis Data Susenas 2004 [Tesis]. Jakarta (ID): Universitas Indonesia.
- Afif, M. N., & Sasana, H. (2019). INDONESIA. *DIPONEGORO JOURNAL OF ECONOMICS*, 1, 88–96.
- Agresti, A. 2002. *Categorical Data Analysis*, Second Edition. New Jersey: John Willey & Sons.
- Azizah N., Amiruddin R., dan Ansariadi (2013). Fakta yang berhubungan dengan perilaku Merokok anak-anak jalanan di kota Makassar.
- Cameron, A.C, dan Trivedi, P.K., (1998). *Regression Analysis of Count Data*. Cambridge University Press
- Dachroni. 1990. *Pem'erintah dan Aspek Ekonomi Pengawasan terhadap Tembakau*. Publikasi Bank Dunia.
- Hamzah, H. (2012). *Analisis Faktor Yang Mempengaruhi Permintaan Rokok Masyarakat di Kota Makassa*.
- Hestiana, Fida. 2013. *Pemodelan Mixture Count Regression Pada Data Zero Inflated dengan Pendekatan Bayesian (Studi Kasus: Jumlah Batang Rokok Yang Dihisap Setiap Hari di Provinsi Sumatera Selatan Tahun 2010)*. Jakarta: Sekolah Tinggi Ilmu Statistik.
- Pontoh, Septiani R dan Faidah, Defi Y. 2015. *Penerapan Hurdle Negative Binomial pada Data Tersensor*. Yogyakarta: UNY.
- Ridout, M., Demetrio, C. G. B. dan Hinde, J. 1998. *Models for Count Data with Many Zero*. International Biometric Conference, Cape Town.

- Santi. 2013. Hubungan Pengetahuan Tentang Rokok dengan Sikap terhadap Bahaya Merokok pada Siswa SMK 1 Batik Surakarta. Fakultas Kedokteran Universitas Muhammadiyah Surakarta. Skripsi. Surakarta.
- Sirait, A. M., Pradono, Y., & Toruan, I. L. (2002). Perilaku merokok di Indonesia. *Buletin Penelitian Kesehatan*, 30(3 Sep).
- Sukirno S. 2008. Mikro Ekonomi Teori Pengantar Edisi Ketiga. Jakarta (ID) : PT Raja Grafindo Persada.
- WHO. (2003). Tobacco Control Country Profile. American Cancer Society, World Health Organization, and International Union Against Cancer, 2003